


```

/gene="YBR0304"
/codon_start=1
/protein_id="CA53679.1"
/db_xref="GI:498751"
/db_xref="GOA:P38218"
/db_xref="InterPro:IPR002589"
/db_xref="SGD:S00000226"
/db_xref="UniProt/Swiss-Prot:P38218"
/translation="MSNITYVKNILPKPSYARILHSQCNQSGWGGGJAYOLALRYP
KAKDVEYCEKXSNLGGKILLPEYNSDLLICLFTSSFGSGSHGKOSLITVYTK
IALDLKTPRAKDKRTSSEDSIGDLNHNHITYPIGEYELMPQINSGLFGVPMKETE
RVLBERSGDMSFTVYOL"
/gene="CAL1"
/complement(7577.11074)
/gene="CAL1"
/complement(7577.11074)
/codon_start=1
/product="chitin synthase 3"
/protein_id="CA53680.1"
/db_xref="GI:498752"
/db_xref="GOA:P29465"
/db_xref="InterPro:IPR001199"
/db_xref="InterPro:IPR004835"
/db_xref="SGD:S00000227"
/db_xref="UniProt/Swiss-Prot:P29465"
/translation="MTGLNGDDDDYTLNLDNDESLRSHSVSGAPHRQSLVPR
ERSLRNPDPNPHFYAOKTQOENHLDVLPSTGVNPNATRRSGSLRSGSVKPSG
RETSYLLDQMTTDEKASVKISDEGADEPDKDQVNPESSTOPIKNSIKPLRK
ERNDTLSFMQMYCYPTTPMAPAPILAFPCMPPKPRQMARREKVALISVILGATYAF
LTFGTIVCSSSKRLKNNVSTEVVINGKAYELDTSSRSGIDVAYEDSDTLVYPM
SDAGKASFLPQVNGNCHNLITPKSNSSI PHDDNNLAMYPCPKLKNODSSKPLPT
VENYAGMNGCHTSKEDDAFYGLSKADYFTWMDGKINSSRNLIYNGVDLIDLDMVL
EKDDVYPVFDLKTNSQGYDLSTLVNGBERKARCLSEILKVGEDSKVGCIA
SDVLLVSLVSLVILKEIILACYFRMTYARQOGAYIDNKTMDGHTDIDBMSNI
OKAPLKEVDPHLPKPKYSKSLGHRKASTPOLIKHSGSKRQPNESVLDIGTMSNSI
LOSSEYRGKTTTTONAKLSNENKAVSRNPSTLLPSSMFKATISPPGSSLI
QSLDSTLIHPDVOQPLDPMFPGPLIHITCVTYSDEBGLRTLLSTLTDIPN
SHLLMNVCDGLIKSGNDKTPPEIAIGMMDDVTPPEDEKPYSAVAVASGRHMA
KIYAGFYKYDSTIPRENOQRVPIITVCGTPEAQGAAPGRGRGDSOILIMSELE
KITPDERMQLFEOLKNIMQITLMAADFEYTLAMDADTKVPDALTVMVAMVDP
LIMGLGCTKILANKASWYTAIOVFYVYISHQAKAPESVFGSVTCLPGCFSMYRIKS
PKSGDGVNVPVLANPDIVERKSDNVNTLHKKNLLLGEDRPLSSMLKTPPKRQVY
VPAACKTIAADKFKVLSQRMRINSTVHNLFEVLIDLCGTFCSMQPVGIELI
GTWVLPALICFTIYVLIIPAVSKPTVITVLAIILGLPGLIVITATRWSTLMWC
VYICALPIMNFVLPYAYMKFDDFSWGDRTTLAGNKKAKQDENEGEFDHSKIKMTWR
EPEFREDIINRKSQDSFVA"
12594.13499
/gene="YBR0308"
12594.13499
/gene="YBR0308"
/codon_start=1
/protein_id="CA53681.1"
/db_xref="GI:498753"
/db_xref="GOA:P38072"
/db_xref="InterPro:IPR003782"
/db_xref="SGD:S00000228"
/db_xref="UniProt/Swiss-Prot:P38072"
/translation="MLNLSRKYACRSIFRANYSIKGLFYNGAVYRGSPSTGCCASD
NKESPAROPDLQGLDEINEPEPIRTFRFQSRMKATIALLLSGTAYAYSRRR
LLETERKADNARAVGSVALGDFPNLIDFNKSPTEBMLKKEFSLTFGFHCDICE
ELDRLTYWISLDDKHKIQPLFISCDPARDPVLYKEVLSFHPAIIIGLTETVQV
KSVCKKXVSTPRDYKPNQIDYLVHDSIFLYIDEGQPIDALGNVDSOGLKXIR
EQLQAYVPEERERRSRKWSYTFN"
complement(13830.15014)
/gene="YBR0309"
complement(13830.15014)
/gene="YBR0309"
/codon_start=1
/protein_id="CA53682.1"
/db_xref="GI:498754"
/db_xref="GOA:P38219"
/db_xref="InterPro:IPR004396"
/db_xref="InterPro:IPR006073"

```

```

/db_xref="SGD:S00000229"
/db_xref="UniProt/Swiss-Prot:P38219"
/translation="MPPKQVEKKVLLGRGNNLKAGIVGLANVKSSTFPOAITRCP
LGNPANYPATIDEKARAVIPSPFDLCIEIYKKTASEVVAHLIVDIDAGLTGASA
GEGGNAPLASHRSYDSIYOVRCPDDEIITHVEDVDPVADLEIINBELKDIEFA
OKALEGAKTIKRGQOSILEVOKKEMDLITRKIIKLESSGQVANNHWSKEVELINS
MFLUTAKPCITLINISESDYIRKKNKHLIRKENWDKXSPDILIIPEVSILBERLSH
SPEDAEELKQLQTSIALPKITITWRQDLDISFTGCPDEVREMTIRGKAPQAG
VIHNDLWMTFLAQWKEDEVEFKDSDAIIKAAKMLKMGKQDVVEDGDIIFRAGAG
KN"
complement(16026.17168)
/gene="YBR0310"
complement(16026.17168)
/gene="YBR0310"
/codon_start=1
/protein_id="CA53683.1"
/db_xref="GI:498755"
/db_xref="GOA:P38071"
/db_xref="SGD:S00000230"
/db_xref="UniProt/Swiss-Prot:P38071"
/translation="MLPTFKRYSMSAHQIPKHFESLIYSTHEVEDCTKVLVKNYTP
KODLSQSIVLKTLAFPIPNSDINLOGVYPSRPEKTYDSTDEPAIAGNEGVFVS
LPSGSKGDKLKGDRVILPQANOGTSMYRVSSSSDLIKVNDLDFSAATVSVNGCT
GEOLYSDYIDNNSNGENWIIIONAGTSVSKITVOYAKKGIKTLISVIRDRDFEYAK
VIEDKYCATKYSISQNDKTPAKEVLSKIIGENARVALANSVGGKSASIGARLEN
NALMTYGMKSQPTLPTSLHIFKGLTSKGYWTERKKNKPNQSKITDITSDTKMNY
GHIIISPREIEFLTWNTTNTTDEQLSLHVKKGIYKGRKKWVLEW"
complement(17574.19151)
/gene="YBR0312"
complement(17574.19151)
/gene="YBR0312"
/codon_start=1
/protein_id="CA53684.1"
/db_xref="GI:498756"
/db_xref="GOA:P38070"
/db_xref="InterPro:IPR000719"
/db_xref="InterPro:IPR000961"
/db_xref="InterPro:IPR002290"
/db_xref="InterPro:IPR008271"
/db_xref="InterPro:IPR011009"
/db_xref="SGD:S00000232"
/db_xref="UniProt/Swiss-Prot:P38070"
/translation="WIFSLDEELHVSDDKNDIKVYSSAIYNDINHEQSSITYE
ESINLHVSANIPPLNGMSPAIRMRKRSASKEPILTPPNRRRSITGSDMRNTN
RSLIPPODISSINIGENELSRNLDHPKPVRYGQAYGVLLKVVNTSKLYAMQAR
KAEILISQATDSKREDEKNDGNNDNDKESKLEETFAERSILSEIHEINIYALF
YSFHNSKLVILLQYIPGGLPYHLKHEGTDDETVVFAAIBISCLARLHTKGVYVR

```

Query Match 97.9%; Score 474.8; DB 15; Length 33117;
 Best Local Similarity 99.6%; Pred. No. 1.8e-80;
 Matches 476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	6	CGGAGGAGGAAGAGACCTCTGCGTGGTGAAGAGTGAATAATTTTACTGTGAT	65
DB	7609	CGGAGGAGGAAGAGACCTCTCTGTCATGACAGATGAATAATTTTACTGTGAT	7550
QY	66	ACTTACAACTGATATATAGTGTGTGTGAACCTTATTTTGAAGATTTTAAACAC	125
DB	7549	ACTTACAACTGATATATAGTGTGTGTGAACCTTATTTTGAAGATTTTAAACAC	7490
QY	126	CTTGAACCTTAAATTAATTAATTTTCTATCTTTTAAAGGACATATTAGCTGG	185
DB	7489	CTTGAACCTTAAATTAATTAATTTTCTATCTTTTAAAGGACATATTAGCTGG	7430
QY	186	CTAAGGCAATTAACAGCTGATATCTGTAACACATGAGCCACTAAATCTCTTAAC	245
DB	7429	CTAAGGCAATTAACAGCTGATATCTGTAACACATGAGCCACTAAATCTCTTAAC	7370
QY	246	GGCTTCTGTCTCTTTCCAAAGGAGCTCCGATATGCCACTATTATCTGTGGCA	305
DB	7369	GGCTTCTGTCTCTTTCCAAAGGAGCTCCGATATGCCACTATTATCTGTGGCA	7310
QY	306	TTTATATTCCTCCTATTGGGATTTTGAATGCGCGTTTAAATATGTCACGATTTG	365

Db 7309 TTTATATCCCTATGGATATGATGAGCGGTTAAATGTCACCGATGAATCTTC 7250

Qy 366 ACTTGTCAGATTTGTCTTTGCTTCTCTTAAAGCTTCATTTATCTAAAGCAAGTTT 425

Db 7249 ACTTGTCAGATTTGTCTTTGCTTCTCTTAAAGCTTCATTTATCTAAAGCAAGTTT 7190

Qy 426 TGTATTAATTCAAAATTAATCTTTGCTTTCTTCATGACTTGAACCTCCAATGATGAGTA 483

Db 7189 TGTATTAATTCAAAATTAATCTTTGCTTTCTTCATGACTTGAACCTCCAATGATGAGTA 7132

RESULT 3
SCYBR022W/c 1089 bp DNA linear PLN 11-AUG-1997

LOCUS S.cerevisiae chromosome II reading frame ORF YBR022w.
ACCESSION Z35891 Y13134
VERSION Z35891.1 GI:536227

KEYWORDS

SOURCE Saccharomyces cerevisiae (baker's yeast)

ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1
Feldmann, H., Aigle, M., Aljinovic, G., Andre, B., Baclet, M.C., Barthe, C., Baur, A., Becam, A.M., Bileau, N., Boles, E., Brandt, T., Breidel, M., Brueckner, M., Bussereau, F., Christiansen, C., Contreras, R., Crouzet, M., Cziepluch, C., Demolis, N., Delaveau, T., Dolignon, F., Domdey, H., Duesterhus, S., Dubois, E., Dujon, B., El Bakoury, M., Enlihan, K.D., Feuerhahn, M., Fiers, W., Fobg, G.M., Fritz, C., Gassenhuber, H., Giansdorff, N., Goffeau, A., Griwell, L.A., de Haan, M., Hein, C., Herbert, C.J., Hollenberg, C.P., Holmstrom, K., Jacq, C., Jacquet, M., Jauniaux, J.C., Joniaux, J.L., Kallioe, T., Kiebau, P., Kirchbach, L., Koelter, P., Korol, S., Liedl, S., Logghe, M., Lohan, A.J.B., Louis, E.J., Li, Z.Y., Maat, M.J., Mallet, L., Manhaupt, G., Messenguy, F., Miosga, T., Molemans, F., Mueller, S., Nasr, F., Obermaier, B., Perez, J., Pierard, A., Pitavandi, B., Pohl, F.M., Pohl, T.M., Potier, S., Prof, M., Purnelle, B., Ramezani Rad, M., Rieger, M., Rose, M., Schaff, G., Schenck, L., Scherens, B., Schwarzi, C., Skala, J., Slonimski, P.P., Smit, P.H.M., Souciet, J.L., Steenma, H.Y., Stucka, R., Uristarazu, A., Van der Aart, O.J., Vana, M.J., Vassart, A., Vetter, I., Vierendeels, F., Vissers, S., Wagner, G., de Wergifosse, P., Wolle, K.H., Zagulski, M., Zimmermann, F.K., Wewes, H.W. and Klein, K. Complete DNA sequence of yeast chromosome II
EMBO J. 13 (24), 5795-5809 (1994)
7813418

TITLE JOURNAL
PUBMED 2 (bases 1 to 1089)
AUTHORS Griwell, L.A., de Haan, M., Maat, M.J. and Smit, P.H.M.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1089)
MIPS.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-1994) Data collected by MIPS on behalf of the European yeast chromosome II sequencing project. MIPS at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152 Martinsried, FRG; E-mail: Mewes@mips.embl.de; org Location/Qualifiers
FEATURES
source 1. 1089
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/chromosome="II"
397. 930
/note="unnamed protein product; ORF YBR022w"
/codon_start=1
/protein_id="CAA84964.1"
/db_xref="GI:536228"
/db_xref="SGD:S0000226"
/db_xref="UniProt/Swiss-Prot:P38218"
/translation="MSNITYVGNILKPKSYAIIHSGNCGNSGGGJAYOLALRYP KAEKDYVECKYGNILGKCIILPSEYNSDLICLPTFSRSGSHGKOSILNWTYK LALDKLTFREAKDKRTSBDSDIGDVLNHIKPIGIEYTLNEMPQINSIGFVPMKETE

ORIGIN RYLFESGDSFTVYQL"

Query Match 92.0%; Score 446.4; DB 15; Length 1089;
Best Local Similarity 99.8%; Pred. No. 1,1e-74;
Matches 447; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 36 TAGACAGTATGAAAATATTTTACTGTGATCTTACAAAGTTGATATATGTTGTGTAA 95

Db 1089 TAGACAGTATGAAAATATTTTACTGTGATCTTACAAAGTTGATATATGTTGTGTAA 1030

Qy 96 TTTATTTATTTAGAGGATTTTAAACACCTTGAACCTAAATATATATATTT 155

Db 1029 CTATTTATTTAGAGGATTTTAAACACCTTGAACCTAAATATATATATTT 970

Qy 156 TCTCATCTTTAAAGGACATATTTAGCGGTGTAAGGACATTTACGTATCTGTAAA 215

Db 969 TCTCATCTTTAAAGGACATATTTAGCGGTGTAAGGACATTTACGTATCTGTAAA 910

Qy 216 ACTCATGTCCGACATTAATTTTCTTACACGCGTTCGTCTTTCCAAAGGACTCCGAA 275

Db 909 ACTCATGTCCGACATTAATTTTCTTACACGCGTTCGTCTTTCCAAAGGACTCCGAA 850

Qy 276 TATGCACTATTTATCTGTGACATTTCCATTTATTTCCCTATTTGGGTATTTGATGTG 335

Db 849 TATGCACTATTTATCTGTGACATTTCCATTTATTTCCCTATTTGGGTATTTGATGTG 790

Qy 336 GCCGTTTAAATGATCACCAGATTTGATCTTCACTTGTGAGTTTGTCTTTGCTTCT 395

Db 789 GCCGTTTAAATGATCACCAGATTTGATCTTCACTTGTGAGTTTGTCTTTGCTTCT 730

Qy 336 AAAGTCTTCAATTTATTTAAGCAAGTTTGTATTAATGAATTTTCTTGTCTTCC 455

Db 729 AAAGTCTTCAATTTATTTAAGCAAGTTTGTATTAATGAATTTTCTTGTCTTCC 670

Qy 456 ATGACTTGAACCTCAATGATGAGTA 483

Db 669 ATGACTTGAACCTCAATGATGAGTA 642

RESULT 4
AY557705/c 534 bp DNA linear PLN 14-MAR-2004

LOCUS Saccharomyces cerevisiae clone FLM00707.01X YBR022w gene, complete cds.
ACCESSION AY557705
VERSION AY557705.1 GI:45269300
KEYWORDS Yeast ORF Project.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 534)
AUTHORS Marischky, G., Rolfe, A., Richardson, A., Kane, M., Bagui, M., Taycher, E., Hu, Y., Vanberg, F., Weger, J., Kramer, J., Moreira, D., Kelley, F., Zuo, D., Raphael, J., Hogle, C., Jepson, D., Williamson, J., Camargo, A., Gonzaga, L., Vasconcelos, A.T., Simpson, A., Kolodner, R., Harlow, E. and Labaer, J.
TITLE Creation of the YFLX clone resource: cloning of Saccharomyces cerevisiae ORFs in the Gateway recombinational cloning system
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 534)
AUTHORS Marischky, G., Rolfe, A., Richardson, A., Kane, M., Bagui, M., Taycher, E., Hu, Y., Vanberg, F., Weger, J., Kramer, J., Moreira, D., Kelley, F., Zuo, D., Raphael, J., Hogle, C., Jepson, D., Williamson, J., Camargo, A., Gonzaga, L., Vasconcelos, A.T., Simpson, A., Kolodner, R., Harlow, E. and Labaer, J.
TITLE Direct Submission
JOURNAL Submitted (17-FEB-2004) Biological Chemistry and Molecular Pharmacology, Harvard Institute of Proteomics, 320 Charles St., Cambridge, MA 02141, USA
COMMENT This clone is part of a collection of Saccharomyces cerevisiae full-length ORF clones generated by the Harvard Institute of

Proteomics. Each CDS has been cloned with its native stop-codon. The CDS has been directionally cloned using the Gateway cloning system into the donor vectors pDONR 201 or pDONR 221. Additional sequences in the clone: 'TCCAGCTGACAC' after the attL1 site and before the 'ATG' (from Research Genetics primers) used to amplify the ORFs, including a Kozak consensus sequence); 'ATCCCCGGAAATGCCATG' after the stop codon and before the attL2 site (from the Research Genetics primers used to amplify the ORFs). Location/Qualifiers

FEATURES

source

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/db_xref="taxon:4932"

/clone="FLH00707.01X"

/lab_host="Escherichia coli DH5alpha T1 resistant"

<1..>534

/product="YBR022W"

1..534

/codon_start=1

/product="YBR022W"

/protein_id="AA556031.1"

/db_xref="GI:45269301"

/translation="MSNITVYKGNLIKPKSVARILHSQNGSGGGIAYOLALRYP

KAKDYVEYCEKYGSLKCLILBSYENSILICLITSSPGSSSHKQSIILATYK

LADLKKTRKADKTRISBDSIGDILNKHITPIGEYKLEMPQINSIGIVPMKETE

RLLEFSGDSFTYVQL"

ORIGIN

Query Match 59.3%; Score 287.4; DB 15; Length 534;

Best Local Similarity 99.7%; Pred. No. 1.6e-44;

Matches 288; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

195 TTACAGCTGATATCTGTAATAAATCTCATGTGCGCACTAAATCTTCTTCAACGCGTTCGT 254
 534 TTACAGCTGATATCTGTAATAAATCTCATGTGCGCACTAAATCTTCTTCAACGCGTTCGT 475
 255 CTCTTTCGAAGGACCTCGAATATGCGCATTTATCTGTGGCATTTCCAAATTTATATTC 314
 474 CTCTTTCGAAGGACCTCGAATATGCGCATTTATCTGTGGCATTTCCAAATTTATATTC 415
 315 CCTATTGGGATTTTGAATGAGCGCGTTTAAATAGTACGCGATTTCACTTCCTGTTG 374
 414 CCTATTGGGATTTTGAATGAGCGCGTTTAAATAGTACGCGATTTCACTTCCTGTTG 355
 375 AGTTTGTCTTGTCTTCTTAAGGCTCTCAATTTATCTTAAAGCAAGTTTGTATAT 434
 354 AGTTTGTCTTGTCTTCTTAAGGCTCTCAATTTATCTTAAAGCAAGTTTGTATAT 295
 435 CAAATATCTTGTCTTCTTCAATGACTTGAACCTCCAAATGATGAGTA 483
 294 CAAATATCTTGTCTTCTTCAATGACTTGAACCTCCAAATGATGATGTA 246

Db

QY 435 CAAATATCTTGTCTTCTTCAATGACTTGAACCTCCAAATGATGAGTA 483
 294 CAAATATCTTGTCTTCTTCAATGACTTGAACCTCCAAATGATGATGTA 246

RESULT 5

YSCPCHSY

LOCUS YSCPCHSY 4345 bp DNA linear PLN 27-APR-1993

DEFINITION S.cerevisiae CSD2 gene (encoding putative chitin synthase), complete cds.

ACCESSION M73697

VERSION M73697.1 GI:172103

KEYWORDS chitin synthase.

SOURCE Saccharomyces cerevisiae (baker's yeast)

ORGANISM Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE 1 (bases 1 to 4345)
 Bulawa, C.E.

AUTHORS CSD2, CSD3, and CSD4, genes required for chitin synthesis in Saccharomyces cerevisiae: the CSD2 gene product is related to chitin synthases and to developmentally regulated proteins in Rhizobium species and Xenopus laevis

JOURNAL Mol. Cell. Biol. 12 (4), 1764-1776 (1992)

PUBMED 1532231

COMMENT Original source text: Saccharomyces cerevisiae DNA.
 FEATURES Location/Qualifiers

source

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/db_xref="taxon:4932"

592..4089

/gene="CSD2"

592..4089

/function="CSD2"

/note="Required for chitin synthesis in vivo and for chitin synthase inactivity in vitro."

/note="Deduced amino acid sequence of CSD2 has limited, but statistically significant similarity to chitin synthases of X. laevis; putative"

protein of X. laevis; putative"

/codon_start=1

/product="chitin synthase"

/protein_id="AAA34844.1"

/db_xref="GI:172104"

/translation="MTGLNGDDPDYILNLNQBESLRSRHSVSGAPHRQSLVR

ERSRLNPDNPHFYAKTOROMHLDVLPSTGVNPAATRRSGLSRSGSRKFS

REDSYLLIDNNITDDKASVKISBGADEDFDGDVDNFBESSTOPIKSLRK

ETNDPLSMQWCVFITFMAAPILABCMGRKQAMREKVALISYILYGAIVAR

LFFGRTKVSCKSLKRLKNEVSTBFVINKAYBLDTSRSGIODVVDSDTLXPM

SDAGKDAFLQNVNNGNHLNLTIPSNSIIPHDNDNLAWFPCKLKNQDSSKPNFT

VENVAGNCHTSEKEDRDAFYGLKSKADVFTFMDGIKSSRLIYVNGVLDLIDLMT

EKDQDVVVFDDKLTSLNIGYDLSLVNNGHERIKASEIIRVGVDSKTVGCLIA

SDVLYVSLVTLVLIIFKIIACFRMTVARKQAYVNDKMTQNDNEDMNNI

OTKPLKQVDPDLRPKYSKSLGKRASTPDLIKSSKMPFQESVYDIDTSMSS

LQSSGSTRGATTTTQNAKLSNENKAVHSNRPSTLPTSSFNKATSSPPGSSLI

QSLSTIIHPDIYVQPLDMPYGPPLHITICFYVSEDESGELTTLIDSLTDYFN

SHKLMMVCDGLKSGNDKTPPEYALCMDDPVPPEVKEYSVAVASGKRNNMA

KITVAGFYKDDSTIPENQOQVPIITIVKCGTAPAGAKNGSKRDSOILNSFLE

KITPBERMTOLEFOLLKNIWOTIGMADFEYELVMDQTVFPDLTHMAVANKDP

LIMGCGEETKIANRAOSVNTAIOVEYVISHOAKAPESVSGVTCILGCSMRVKS

PKGSDGVVPLANPDIYERYSDVNTTLHKNNLLLEDEPLSLMKITPPKQOV

VPKACQKIADPKRVLISQRRMNINSVHNLFEVLRLDLCGTFCSFMOVPTGLEI

GTMVLPALICETIYVLIIFAIYSKPTVTLVLVLIIGLPLIVITVITRYSYLMWC

VYICALPIWNLVPSYAWKFPDPSWGTTRTLAGNKAQDENBGEFDSKIKMRTWR

EPPERDLINRKBSDSLVA"

ORIGIN

Query Match 58.6%; Score 284.2; DB 15; Length 4345;

Best Local Similarity 99.0%; Pred. No. 3.4e-44;

Matches 286; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 CGAAGGAGGAAGGACCTCTGTCGTRGACAGTATGAAATATTTTACTGTGAT 65
 DB 4057 CGAAGGAGGAAGGACCTCTGTCGTRGACAGTATGAAATATTTTACTGTGAT 4116
 QY 66 ACTTACAGTTGATATATGATGTTGTGTAATTTATTTAGAGATTTTAAACAC 125
 DB 4117 ACTTACAGTTGATATATGATGTTGTGTAATTTATTTAGAGATTTTAAACAC 4176
 QY 126 CTTAGACATAAATCTTAATAATAATATTTCTATCTTTAAAGCACATTTACGTG 185
 DB 4177 CTTAGACATAAATCTTAATAATAATAATATTTCTATCTTTAAAGCACATTTACGTG 4236
 QY 186 CTAAGGCAATTCAGCTGATATATCTGTAATAAATCTATGCGCACTAAATTTCTTAAAC 245
 DB 4237 TTAAGGCAATTCAGCTGATATATCTGTAATAAATCTATGCGCACTAAATTTCTTAAAC 4296
 QY 246 GCGTCTGCTCTTTCGAAGGACCTCGAATATGCGCATTTATCTGT 294
 DB 4297 GCGTCTGCTCTTTCGAAGGACCTCGAATATGCGCATTTATCTGT 4345

RESULT 6

SCYBR023C

LOCUS SCYBR023C/c 5176 bp DNA linear PLN 11-AUG-1997

DEFINITION S.cerevisiae chromosome II reading frame ORF YBR023C.

ACCESSION Z35892 Y13134

VERSION 235892.1 GI:536229
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
JOURNAL
FEATURES
source
gene
CDS
Saccharomyces cerevisiae (baker's yeast)
Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
1
Feldmann, H., Aigle, M., Aljinovic, G., Andre, B., Baclot, M. C.,
Bartel, C., Baur, A., Becam, A. M., Bileau, N., Boles, E., Brandt, T.,
Brendel, M., Bruckner, M., Busecraeu, F., Christiansen, C.,
Contreas, R., Crouzet, M., Czaplich, C., Demolis, N., Delaveau, T.,
Dognon, F., Domery, H., Duesterhus, S., Dubois, E., Dujon, B., El
Bakoury, M., Enliam, K. D., Feuermann, M., Fiers, W., Fobo, G. M.,
Fritz, C., Gassenhuber, H., Glandsdorf, N., Goffeau, A., Grivell, L. A.,
de Haan, M., Hein, C., Herbert, C. J., Hollenberg, C. P., Holmstrom, K.,
Jaceq, C., Jacques, M., Jaumaux, J. C., Jomiaux, J. L., Kallioe, T.,
Klesau, P., Kirschner, L., Koetter, P., Korol, S., Lisdel, S., Logghe, M.,
Lohan, A. J. E., Louis, E. J., Li, Z. Y., Maat, M. J., Mallet, L.,
Manhaupt, G., Messenguy, F., Miesse, T., Molemans, F., Mueller, S.,
Naer, F., Obermaier, B., Pera, J., Pierard, A., Piravandi, B., Ramezani
Pohl, F., Pohl, T. M., Potier, S., Proft, M., Purnelle, B., Ramezani
Rad, M., Rieger, M., Rose, M., Schaaf, G., Schaefer, I.,
Schierens, B., Schwarzlose, C., Skala, J., Slonimski, P. P.,
Smits, P. H. M., Souciet, J. L., Steensma, H. Y., Stucka, R.,
Urestarazu, A., Van der Aart, Q. J., Van Dyck, L., Vassarotti, A.,
Vetter, I., Vierendeels, F., Vissers, S., Wagner, G., de Wergifosse, P.,
Wolfe, K. H., Zagulski, M., Zimmermann, P. K., Mewes, H. W. and Kline, K.
Complete DNA sequence of yeast chromosome II
EMBO J. 13 (24), 5795-5809 (1994)
7813418
2 (bases 1 to 5176)
Grivell, L. A., de Haan, M., Maat, M. J. and Smits, P. H. M.
Unpublished
3 (bases 1 to 5176)
MIPS.
Direct Submission
Submitted (30-AUG-1994) Data collected by MIPS on behalf of the
European yeast chromosome II sequencing project. MIPS at the
Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
Martinsried, FRG; E-mail: Mewes@mips.emblnet.org
Location/Qualifiers
1. 5176
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
/chromosome="II"
complement(160..3657)
/gene="CHS3"
complement(160..3657)
/gene="CHS3"
/note="ORF YBR023c"
/codon_start=1
/protein_id="CA84965.1"
/db_xref="GI:536230"
/db_xref="GGA:P29465"
/db_xref="SGD:S0000227"
/db_xref="UniProt/Swiss-Prot:P29465"
/translation="MTGLNDPDDPDYILNNOBESILASRRHVSQSGAHPHRSIVRP
ERSLNDPNDPPIYAKTOEONHNDVLPSTGVFNATREGSRKSGVSRKSG
RETDYLDQMTNTDKKASVKISDEGVADEPKDDVDNDEFSSTQPIKSLPK
ETNDYLFQMGYCYFITFMAFAPILAFCCGPKPKDQAMREKALISVILYGLIYAF
LTFGTFKTVQSSSKRLKNNSESTEFVINGKAYELDTSRSGIIOVEVSDLYGMP
SDAGKASPLRQVNGNCHNLITPKNSISPHDDNNILAMPCKLKNQDSKSPPT
VENYAGNCHTSKEDRDAFYGLSKADYVTPMGIKSSSNLIVNGDYLIDLML
EKQVDYPPVFDLKTSLQGYDLSLVLSGHRKAKLAKLSIITKGEVDSKTVGCL
SDVLYVSLVFIISVLIKPIIACYFMTAARQAGATVINKMTRHTNDIEDWNNI
OTAPLKEVDPHLPKPKYSKSLGKKASFFDLKKHSSKMFQNESVFDLDTSMSS
LQSSGYRGTMTQNLQAMKLSNENKALSRNPSLTPTSFMNKAISGPPGSSLI
OSIDSTIIPDIQOPLDPMYGPPIIHTICVTCSDEBGRATLIDSLTDYDN
SHKLNAVCDGLIKSGNDKTTPIALGMMDDVTPPDYKPSYVAVAGSRHMA
KIYAGFYKDDSTLIPENQORVPLITIVKCGTAPGGAARNGRGRDSQIILMSLE
KITFERMTQLEPOLINQOITGLMADFEYELVADADTKVPDALTTHVAVMADP
LIMLCGETKIANKAGSWTALQVFEFYIISHQAKAKESVSGVTCILPGCFSMYRIKS

PKSGDGYWPIVLANPDIVERYSNVTNTLHKNNLLIGEDRPLSSMLTKPPKROVF
VPAACKTAPDKFKVLISQKRWINSTVHNPFLVILNDLCTGTPERSQNPVIGIELI
GIMVLPALCFITVYIIFAIIVSKPVPVITVLALITLGLLIVTATRWSTYLMNC
VYICALPINWFLVPSYAWKFDPSWGDRTIAGNKKAQDENEGBFSDSKJIKRITWR
EPEERILNRKESDSFVA"
complement(4476..4486)
/note="ARS-consensus"
ORIGIN
misc_feature
Query Match 39.3%; Score 190.4; DB 15; Length 5176;
Best Local Similarity 99.5%; Pred. No. 1.7e-26;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
6 CGGAGGAGGAGAAAGTGACTCTCTGTTGGCTAGACAGTAAAGAAATTTTCTGAT 65
192 CGGAGGAGGAGAAAGTGACTCTCTGTTGGCTAGACAGTAAAGAAATTTTCTGAT 133
QY 66 ACTTACAACTTGAATATATGTTGGTGTGTAATTAATTTGAGGATTTTAAACACAC 125
Db 132 ACTTACAACTTGAATATATGTTGGTGTGTAATTAATTTGAGGATTTTAAACACAC 73
QY 126 CTTAGAACTAAACTTAAATTAATTAATTTCTATCTTTAAAGCACATATTACGTGG 185
Db 72 CTTAGAACTAAACTTAAATTAATTAATTTCTATCTTTAAAGCACATATTACGTGG 13
QY 186 CTAAGGCAATTA 197
Db 12 CTAAGGCAATTA 1
RESULT 7
CR380948_4
WPCOMMENT
Sequence split into 5 fragments LOCUS CR380948 Accession CR380948
Fragment Name Begin End
CR380948_0 1 110000
CR380948_1 100001 210000
CR380948_2 200001 310000
CR380948_3 300001 410000
CR380948_4 400001 502101
Continuation (5 of 5) of CR380948 from base 400001 (CR380948 Candida glabrata strain CBS
Query Match 13.0%; Score 63.2; DB 15; Length 102101;
Best Local Similarity 53.4%; Pred. No. 0.0075;
Matches 159; Conservative 0; Mismatches 133; Indels 6; Gaps 1;
QY 192 CAATTAACGCTGATATATGTAATTAATTCATGCGCACATAATTTCTTAACACGCGTTC 251
Db 31268 CATTGAGATTCAATACAGTAATTAATTCATGATGTCATACCTTTTAATTCCTC 31327
QY 252 TGTCTCTTCCAAAGGACTCCGAATATGCACTATTATCTGTGGCAATTTCAATTATA 311
Db 31328 GGTTCCTGGCCAAAGGACTCCGAATATACCGTATTTGATTTTGGCATCTCGAGCTTGA 31387
QY 312 TTTCCCTATGCGATATTGATGTGCGCGCTTAAATGATCAGGATTTGATCTTCACTTGT 371
Db 31388 ATCTTAAGTACTCCGCTCTTCAAGTTCAATGTTTCTCAATTTCTTAACAGATTGC 31447
QY 372 TCGA-----GTTTGTCTTTGCTTCTCTAAAGCTTCAATTTATCTAAAGCAAGTTT 425
Db 31448 GTTATCAATTTACTCTGTTCTTTTCCCAACAATCATCAAAATGATGATGAGCTTGTC 31507
QY 426 TGTATTAATCAAAATACCTTGTCTTCTCCATGACTGAACCTCCAAATGATGAGGTA 483
Db 31508 AGTATTAATCAACATATCTAGTGCGCGTGCATGCGCTCCGAATATTTGATGTA 31565
RESULT 8
CR382399 348174 bp DNA linear INV 17-APR-2005
LOCUS CR382399 fa1c1parum chromosome 6, complete sequence; segment 2/5.
DEFINITION CR382399 A1844505
ACCESSION CR382399.1 GI:46361038
VERSION

KEYWORDS HTG.
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 348174)
AUTHORS Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,D., Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K., Buckee,C.O., Butrows,C., Cherevach,I., Chillingworth,C., Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corton,C., Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J., Felwell,T., Goble,A., Goodhead,I., Gilliam,R., Hamlin,N., Hane,Z., Harper,D., Hauser,H., Hornsbey,T., Holroyd,S., Horrocks,P., Humphray,S., Jagers,K., James,K.D., Johnson,D., Kerriourou,A., Knights,A., Konfortov,B., Kyes,S., Laike,N., Lawton,D., Lennard,N., Line,A., Maddison,M., McLean,J., Mooney,P., Moulon,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A., Rabinowitch,E., Rajandream,M.A., Rutter,S., Sutherland,K.M., Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R., Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,L., Whitehead,S., Woodward,J., Sulston,J.E., Craig,A., Newbold,C. and Barrett,B.G.
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
PUBMED 12368867
REFERENCE 2 (bases 1 to 348174)
AUTHORS Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K., Berriman,M., Pain,A., Hall,N., Atkin,R., Chillingworth,C., Doggett,J., Ormond,D., Sanders,M., Hayes,R., Hall,S., Quail,M. and Barrett,B.G.
TITLE Unpublished
JOURNAL 3 (bases 1 to 348174)
AUTHORS Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K., Berriman,M., Pain,A., Hall,N., Atkin,R., Chillingworth,C., Doggett,J., Ormond,D., Sanders,M., Hayes,R., Hall,S., Quail,M. and Barrett,B.G.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2002) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
REFERENCE 4 (bases 1 to 348174)
AUTHORS Cherevach,I., Davis,P., Goodhead,I., Stevens,K., Mungall,K., Berriman,M., Pain,A., Hall,N., Atkin,R., Chillingworth,C., Chillingworth,C., Doggett,J., Ormond,D., Sanders,M., Hayes,R., Hall,S., Quail,M. and Barrett,B.G.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-2004) P.falciparum Genome Sequencing Consortium, The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
source 1. 348174
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="6"
complement(847..1143)
/locus_tag="PF0415C"
complement(847..1143)
/locus_tag="PF0415C"
/old_locus_tag="MAL6P1.87"
/codon_start=1
/product="hypothetical protein, conserved"
/protein_id="CAG25326.1"
/db_xref="GI:46361039"
/db_xref="UniProt/TREMBL:Q6LF12"
/translation="MNFSKILPLKXHQTIIRNVSLSKNGNHMFKEAATSTENKQ EOKQVENVYVANCCKGKIRIPMRVFGILFLMAFVPTQSLYETNKYKEVS"
complement(join(2408..3006,3295..3403))
/locus_tag="PF0420C"
complement(join(2408..3006,3295..3403))
/locus_tag="PF0420C"

gene
/old_locus_tag="MAL6P1.88"
/EC_number="3.4.25.1"
/codon_start=1
/product="proteaseome subunit alpha type 2, putative"
/protein_id="CAG25327.1"
/db_xref="GI:46361040"
/db_xref="GOA:Q6LFI1"
/db_xref="InterPro:IPR001353"
/db_xref="UniProt/TREMBL:Q6LF11"
/translation="MADGEYSFSLTSPTEKLVQIEYALNRVSSSPALGIRAKGV IAREKSPNLIENSIFPKIQOSEHGIYVAGAPGDFVLRLKARKEAARYLQYGV SEILKELVXTASIVQEPOTGVRPFGSLILQGVYGVTHLYQIDPSCTYMAA TCVGKQVNMNSFLEKRNKQIEEDAIHTAIITLKESYEGVLNKEKNIIEGVAIDNKP FKILQNEIKYLLIE"
join(5857..5887,6084..6904)
/locus_tag="PF0425W"
join(5857..5887,6084..6904)
/locus_tag="PF0425W"
/old_locus_tag="MAL6P1.89"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAG25328.1"
/db_xref="GI:46361041"
/db_xref="UniProt/TREMBL:Q6LF10"
/translation="MEDSSLEKSIKGEDELESEVNEQNEBSENESEVNESE NESEVNEVNEENKNEKNKNQNMVQENNNNNNNNNKNDKLNKSFHNNEL KNNLNNN NNKSNKSFKNNEHMMIDQKNSKHEKRNRYDNNLVFPQSDKRNRLFQGNKKEFTFP SLSPDVILPEKHEHNTIMVHFNEPEKGYKNNNGLKPSBEGYLLYPHISHNSIAPKIK KKKRRC"
join(9019..10528,10710..10831)
/locus_tag="PF0430W"
join(9019..10528,10710..10831)
/locus_tag="PF0430W"
/old_locus_tag="MAL6P1.90"
/codon_start=1
/product="chaperone, putative"
/protein_id="CAG25329.1"
/db_xref="GI:46361042"
/db_xref="GOA:Q6LFH9"
/db_xref="InterPro:IPR001844"
/db_xref="InterPro:IPR002194"
/db_xref="InterPro:IPR002423"
/db_xref="UniProt/TREMBL:Q6LFH9"
/translation="MSHILNKKKADSLSTVNLMTNNMNSKMYEIIKSNLPGKSYK MLYSSGAIKTKDGNVLYNEMMIQHPASVLSKSIDETLGDGSSNIVATSLI VLSKYLIVENIHRIITQGGDIKAGIKPVIENPNDEKLIVNCKTCI RTKPLICADPLADDLVESIKIITYRPTQIDLHMEIMDIDRNNNSINTKLVRGWYLDH GCRHPNMNKLTGCFILVNLVLSIEKSEVSFSVYSAEDRDKLVSEKFTDQVK KTIIEKILVKKREKTEINIEYFAVFNQKIDPISLDLAKENIMALKRIRMLERI VILCGGPNVNDVLTBEDVYAGLYVEISINDKYPFIEVQNPKSCITFIQANDY TIKQIKQAIRDGLSKIKVVIDKCVLSAGSFEIMAYKIKDEKKIKGKQKFAIDY ANSLINIPKVLINSGLDIHOTLFNVIDKNVDESEPLGLDLTGEPIIHLKGIYDN YCVKKEIISITATISQILLVDEIIRACKSGEER"
13263..14507
/locus_tag="PF0435W"
13263..14507
/locus_tag="PF0435W"
/old_locus_tag="PF0435W"
/old_locus_tag="MAL6P1.91"
/EC_number="2.6.1.13"
/codon_start=1
/product="ornithine aminotransferase"
/protein_id="CAG25330.1"
/db_xref="GI:46361043"
/db_xref="GOA:Q6LFH8"
/db_xref="InterPro:IPR005814"
/db_xref="InterPro:IPR010164"
/db_xref="UniProt/TREMBL:Q6LFH8"
/translation="MDFPYKELSSQDYNNNMLTYGAHNYDPIPVYLKRGKGVFYDIE DRYVDYFLSASVSVQGHCHDIIINAMINQAKLTI CSRAPFSLSIGYCEBYLNLPG YDVLMMNTGAASETAYKIKCRKGYEYKIPENSAKLIYCNNNPSGLGCLCVSSTD KKKCNRRGPPVNPFLKVPYDDELALRELEODPNVCAPIVEVQEGAGVTPSDSYFPG VASLCKRTYNVLVADDEVDTGIGRTGKLCTHHYGVKPVYIILGKALSGSHPIAIIIA

RESULT 10
AC079954 158785 bp DNA linear PRI 19-MAR-2003
LOCUS Homo sapiens 12 BAC RP11-179A1 (Roswell Park Cancer Institute Human
DEFINITION BAC library) complete sequence.
ACCESSION AC079954.18 GI:13899365
VERSION AC079954.18
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukayocva; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 158785)
Muzny,D.M., Adams,C., Adio-Oduola,B., All-ouman,F.R., Allen,C.,
Alshbrook,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
Burich,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emertling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,U., Harris,C., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,U., Hernandez,O.,
Hodgson,A., Hognes,M., Hollway,C., Hollins,B., Homsl,F.,
Howard,S., Huber,J., Huiyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,U., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W.,
Loulsegad,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,B., Massey,E., Mawhinney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,N., Nguyen,N., Nguyen,N., Nickerson,E., Nwokweto,S.,
Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shin,C.,
Shoohartari,N., Sisson,I., Sodergren,E., Sonalke,T., Spatke,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Umanu,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamsen,A.,
Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorrilla,S., Zuchterlapati,R. and Gibbs,R.

TITLE Unpublished
JOURNAL 2 (bases 1 to 158785)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (20-SEP-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 158785)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 158785)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission

JOURNAL
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 158785)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (25-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 158785)
REFERENCE
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (19-MAR-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 1, 2001 this sequence version replaced gi:13877202.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

COMMENT
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards. Estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

Summary Statistics		
Contig length:	158785	
Phrap values in estimate:	158572	
Average error rate (BCM-Phrap estimate):	4.46022e-06	
Fraction of Phrap values less than 40 :	0.00425044	
Number of consensus changing edits:	30	
Number of N's in consensus :	0	

----- Consensus changing edits -----

Position	Original+Context	Edited+Context
10464	ctctctctc(n)ctnctctctc	cttctctctc(t)ctctctctc
10467	ctctctctc(n)ctnctctctc	ctctctctc(c)ctctctctc
10469	ctctctctc(n)ctctctctc	ctctctctc(c)ctctctctc
10491	ctctctctc(n)ctctctctc	ctctctctc(c)ctctctctc
10511	ctctctctc(n)ctctctctc	ctctctctc(c)ctctctctc
13930	taccaccatc(n)ctnctnctcag	taccaccatc(c)ctcgcacacag

```

13931 acccactcn (n) tmmncaaga acccactcc (c) tgcacaaga
13933 ccactcntr (n) nmcaagaac ccactccct (g) caacaagaac
13934 ccactcntr (n) nmcaagaac ccactccct (g) caacaagaac
13935 cactcntrn (n) ncaagaact cactccctg (a) acaagaact
13936 actcntrn (n) caagaactc actccctg (a) acaagaact
13937 acaaatc (n) ggaatagat acaaatc (a) tgcataagat
13938 attngcaa (n) tagatccaa attngcaa (a) tagatccaa
13939 acaatrat (n) cagatgata acaatrat (a) cagatgata
13940 tatgactaag (n) gaattacc tatgactaag (c) gaattacc
13941 gacacaaag (n) gaattacc gacacaaag (c) gaattacc
13942 aagtcctca (n) aagaagagc aagtcctca (g) aagaagagc
13943 acccactcc (n) aaaaaaaa acccactcc (c) aaaaaaaa
13944 gtaattcag (n) tctttttt gtaattcag (c) tctttttt
13945 ctagncttr (n) tttttttt ctagncttr (c) tttttttt
13946 accgaccca (n) ccaatttca accgaccca (c) ccaatttca
13947 ctatattcg (n) aacctgttaa ctatattcg (c) aacctgttaa
13948 aagtgagga (n) gctttttaa aagtgagga (g) gctttttaa
13949 cgttgcmaa (n) cgttcagag cgttgcmaa (g) cgttcagag
13950 acccactatg (n) tggacaagt acccactatg (g) tggacaagt
13951 aacgaaga (n) naacaagct aacgaaga (a) naacaagct
13952 acagcaagan (n) aacaagct acagcaagan (a) aacaagct
13953 tggagtcag (n) ggcacatct tggagtcag (c) ggcacatct
13954 gtaattcag (n) agagacgag gtaattcag (c) agagacgag
13955 ctctcagaga (a) t. . . . . ctctcagaga (t) t. . . . .

```

----- Distribution of Quality < 40 Bases -----

```

#      500
450    *
400    *
350    *
300    *
250    *
200    *
150    *
100    *
50     *
0      *

```

Phrap Value Range

Version: 1.01 gxf.

Location/Qualifiers

FEATURES

```

source
1. .158785
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-179A1"
1. .1977
/notes="overlaps bases 105124..107100 of clone AC117377"
/function="clone overlap"
misc_feature
1375..1611
misc_feature
1375..1611

```

Query Match 10.6%; Score 51.4; DB 8; Length 158785;
 Best Local Similarity 49.6%; Pred. No. 1.1;
 Matches 186; Conservative 0; Mismatches 186; Indels 3; Gaps 2;

```

QY 43 TATGAAGATTTTACTGATGATCTACAGAGTGTATATGTTGCTGTAATTT 102
DB 135259 TATGACTTTTCTCTTCCCTTGAATATTAATAATTTAAATTCCTTTCTTAAATTAATTT 135318
QY 103 ATTGAAGAGTATTTTACACACCTTAGAATTAATAATAATTAATTTCTCTAT 162
DB 135319 ATTTCAGAAATTTTGCRAATTCAGTAATCTGTAATTTATTTATTCAGGACATTTTAC 135378
QY 163 CTTTAAAGGACATATTTACGCTGCTAAAGCAATTACAGCTGATATCTGTAATCTCATG 222
DB 135379 C-TCATTAACAAACATGAGAGTTTCTTTATTTGATGCTTTACTTACGCTACACTTT 135437

```

```

QY 223 TCOCACATTAATTTCTTAAACAGCGCTTGTCTCTTTCCAGGAATATGCCA 282
DB 135438 TCCTTTCAAAATTTCTATCATATGATGATGATC--TCATGATCTTTCCCTGTGTTT 135495
QY 283 CTATTTATCTGAGCATTTCCATTTTATATCCCATTTGGATTTGATGAGCGGTTT 342
DB 135496 TATCTTTTCTCTGTATTTTATTTTATATATGATTTCTGTCTATTAATGCTTAAATAT 135555
QY 343 AAATAGCACGATGATGATCTTCACTGTCAGATTGTTGTTTGGTTCTTAAAGTTC 402
DB 135556 TCATCTTACTGACTTGGGTTTCTCATGATCAGTATTTTTCCTTTTACAGTTT 135615
QY 403 TTCAATTTATCTMAA 417
DB 135616 TGAATTTGATATGAA 135630

```

RESULT 11
 AC123054/c 174874 bp DNA linear ROD 05-NOV-2003
 LOCUS Mus musculus BAC clone RP24-220N11 from 5, complete sequence.
 DEFINITION AC123054
 ACCESSION AC123054.2 GI:22213530
 VERSION HTG.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murine; Mus.

REFERENCE
 AUTHORS Nguyen,C., Bielecki,L. and Creason,K.
 TITLE The sequence of Mus musculus BAC clone RP24-220N11
 JOURNAL Unpublished (2001)
 REFERENCE
 AUTHORS Wilson,R.
 TITLE Sequencing of Mus musculus
 JOURNAL Unpublished (2001)
 REFERENCE
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE 3 (bases 1 to 174874)
 JOURNAL Submitted (27-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE
 AUTHORS McPherson,J.D. and Waterston,R.H.
 TITLE 4 (bases 1 to 174874)
 JOURNAL Submitted (29-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE
 AUTHORS Wilson,R.
 TITLE Direct Submission
 JOURNAL Submitted (13-AUG-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE
 AUTHORS Wilson,R.
 TITLE Direct Submission
 JOURNAL Submitted (05-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Aug 13, 2002 this sequence version replaced gi:21217626.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@wustl.edu
 ----- Summary Statistics
 Center project name: M_BB0220N11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
<http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCT-24 BAC Library has been constructed by Pieter de Jong and
coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen
and/or brain genomic DNA. The clone and detailed information can be
obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
source
Location/Qualifiers

```
1..174874
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/map="5"
/clone="RP24-220N11"
/clone_1b="RPCT-24"
1..481
/rpt_family="L1"
repeat_region
1533..1639
/rpt_family="L1"
repeat_region
1635..2294
/rpt_family="L1"
repeat_region
2312..2541
/rpt_family="L1"
repeat_region
2551..3982
/rpt_family="L1"
repeat_region
3980..5313
/rpt_family="L1"
repeat_region
5404..7065
/rpt_family="L1"
repeat_region
7314..8273
/rpt_family="L1"
repeat_region
8831..8883
/rpt_family="ERV1"
repeat_region
9597..9729
/rpt_family="L1"
repeat_region
10923..11059
/rpt_family="Alu"
repeat_region
11394..11592
/rpt_family="MER1_type"
repeat_region
11914..12004
/rpt_family="L1"
repeat_region
12002..13641
/rpt_family="L1"
repeat_region
13517..15573
/rpt_family="ERV1"
repeat_region
15824..15915
/rpt_family="Alu"
repeat_region
16175..16378
/rpt_family="L1"
repeat_region
16379..16526
/rpt_family="Alu"
repeat_region
16527..16808
/rpt_family="L1"
repeat_region
16805..17645
/rpt_family="L1"
repeat_region
17893..17982
/rpt_family="MER2_type"
repeat_region
```

```
repeat_region 18413..18629
/rpt_family="L1"
repeat_region 18689..19125
/rpt_family="L1"
repeat_region 19126..19511
/rpt_family="MALR"
repeat_region 19512..19532
/rpt_family="L1"
repeat_region 24546..24780
/rpt_family="L1"
repeat_region 24817..25216
/rpt_family="L1"
repeat_region 25285..25831
/rpt_family="L1"
repeat_region 27096..27471
/rpt_family="MALR"
repeat_region 28325..28446
/rpt_family="B4"
repeat_region 28357..28679
/rpt_family="MALR"
repeat_region 28844..29149
/rpt_family="MALR"
repeat_region 29164..29377
/rpt_family="L1"
repeat_region 30082..30643
/rpt_family="L1"
repeat_region 30806..31168
/rpt_family="MALR"
repeat_region 31398..31799
/rpt_family="L1"
repeat_region 34685..34821
/rpt_family="MALR"
repeat_region 34714..34835
/rpt_family="Alu"
repeat_region 34871..35975
/rpt_family="L1"
repeat_region 35981..36481
/rpt_family="L1"
repeat_region 36468..36857
/rpt_family="L1"
repeat_region 36863..37796
/rpt_family="L1"
repeat_region 37826..38542
/rpt_family="L1"
repeat_region 38875..38977
/rpt_family="L1"
repeat_region 38992..39372
/rpt_family="L1"
repeat_region 39377..39564
/rpt_family="ERVK"
repeat_region 39466..39660
/rpt_family="B4"
repeat_region 39664..39985
/rpt_family="L1"
repeat_region 41371..41452
/rpt_family="L1"
repeat_region 41714..41854
/rpt_family="Alu"
repeat_region 42684..42717
/rpt_family="L1"
repeat_region 42713..43611
/rpt_family="L1"
repeat_region 44642..45115
/rpt_family="L1"
repeat_region 45116..45922
/rpt_family="L1"
repeat_region 46124..46424
/rpt_family="L1"
repeat_region 47263..47408
/rpt_family="Alu"
repeat_region 48705..49182
/rpt_family="L1"
repeat_region 49585..49839
```



```
repeat_region 13838..14223 /rpt_family="MaLR"
repeat_region 14224..14660 /rpt_family="L1"
repeat_region 14720..14936 /rpt_family="L1"
repeat_region 15367..15456 /rpt_family="MER2_type"
repeat_region 15704..16544 /rpt_family="L1"
repeat_region 16541..16822 /rpt_family="L1"
repeat_region 16823..16970 /rpt_family="Alu"
repeat_region 16971..17174 /rpt_family="L1"
repeat_region 17434..17525 /rpt_family="Alu"
repeat_region 17776..17832 /rpt_family="ERV1"
repeat_region 19708..21347 /rpt_family="L1"
repeat_region 21345..21435 /rpt_family="L1"
repeat_region 21757..21955 /rpt_family="MER1_type"
repeat_region 22290..22426 /rpt_family="Alu"
repeat_region 23620..23752 /rpt_family="L1"
repeat_region 24455..24518 /rpt_family="ERV1"
repeat_region 25076..26035 /rpt_family="L1"
repeat_region 26284..27945 /rpt_family="L1"
repeat_region 28036..29369 /rpt_family="L1"
repeat_region 29367..30798 /rpt_family="L1"
repeat_region 30808..31037 /rpt_family="L1"
repeat_region 31055..31714 /rpt_family="L1"
repeat_region 31710..31816 /rpt_family="L1"
repeat_region 32868..34208 /rpt_family="L1"
repeat_region 34267..34341 /rpt_family="MaLR"
repeat_region 35856..35925 /rpt_family="L1"
repeat_region 36371..37274 /rpt_family="MaLR"
repeat_region 37824..37906 /rpt_family="ERV1"
repeat_region 38023..38355 /rpt_family="ERV1"
repeat_region 38655..39782 /rpt_family="L1"
repeat_region 39782..39923 /rpt_family="L1"
repeat_region 39924..40316 /rpt_family="MaLR"
repeat_region 40317..41412 /rpt_family="MaLR"
repeat_region 41413..41806 /rpt_family="MaLR"
repeat_region 41807..42069 /rpt_family="L1"
repeat_region 45555..45675 /rpt_family="L1"
repeat_region 48259..48354 /rpt_family="L1"
```

```
repeat_region 48777..49054 /rpt_family="ERV1"
repeat_region 49190..50095 /rpt_family="L1"
repeat_region 50111..50185 /rpt_family="L1"
repeat_region 50178..50811 /rpt_family="L1"
repeat_region 50812..51133 /rpt_family="L1"
repeat_region 51765..51952 /rpt_family="B4"
repeat_region 52274..52574 /rpt_family="B4"

Query Match 10.6%; Score 51.4; DB 9; length 206926;
Best local similarity 48.2%; Pred. No. 1;
Matches 145; Conservative 0; Mismatches 156; Indels 0; Gaps 0;

QY 17 AAGTACCTCTGTTGCTGAGACATGAAATATTTTACTGTGATCTTACAGTT 76
DB 10973 AAACACATTTTATTTTATATATATTTTACATGACTTGTGACATTTACATTT 11032
QY 77 GATATATGTTGTGTCTACTTATTTATTTGAGAGTATTTTACACACCTTAGACTTA 136
DB 11033 CATATGCAATFACATGAGAACTTAAATGATATTTTCTATTCACGACCTTAACCTCA 11092
QY 137 AACTTAAATTAATTAATTTCTTATCTTTAAAGCACATATTACGTGCTAAGCAATT 196
DB 11093 CAGTTATATTTTATTTCTTTGATATTTAGTATATGAAATATTAATTTTACTTTT 11152
QY 197 ACAGCTGATATACGTGTAACATCATGTCGCCATTAATTTCTTCAACACGCTTCTGT 256
DB 11153 TAAATTAATATTTTCTGAAATTTCTGCACTGTGTATTTGATTCACATCACTTTTACCCT 11212
QY 257 CTTTCAAGGACCTCGGATATGCACTTTATCTGTGGCACTTCCATTTATATTTCC 316
DB 11213 CCTTCAATTTTAACTCCCTCTTCAACCTTTCATTTGCCATTTACAGTTTCATCA 11272
QY 317 C 317
DB 11273 C 11273

RESULT 13
AC087277 196859 bp DNA linear PRI 24-APR-2002
LOCUS AC087277
DEFINITION Homo sapiens chromosome 11, clone RP11-514F3, complete sequence.
ACCESSION AC087277
VERSION AC087277.11 GI:20147901
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 196859)
REFERENCE 1 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-514F3
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 196859)
Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bassett,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,P., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagob,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., Labocque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Menus,L., Minova,T., Munga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (23-DEC-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 196859)

O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Sounez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Travers, M., Travers, M., Trivis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (14-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 196859)

Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boughalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fairo, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gargyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Johnson, R., Jones, C., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Melrim, J., Menus, L., Milnova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Trivis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (24-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 14, 2002 this sequence version replaced gi:19683162.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

FEATURES

source

----- Project Information
Center project name: L11835
Center clone name: 514_F_3

Location/Qualifiers

1. 196859
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-514F3"
/clone_lib="RP11-11 Human Male BAC"
complement(1362..1667)
/rpt_family="AluSg"
complement(1696..2027)
/rpt_family="MLTIF1"
2032..2271
/rpt_family="AluYb"
complement(2441..2542)
/rpt_family="L2"
2700..2894
/rpt_family="L1MC4a"
complement(2895..3268)
/rpt_family="MLT1A1"
3276..3995
/rpt_family="L1MC4a"
complement(3996..4429)
/rpt_family="AluSg"
4300..4355
/rpt_family="L1MC4a"
4703..4723
/rpt_family="AT-rich"
4808..4864
/rpt_family="L2"
complement(5756..5916)
/rpt_family="MER58A"
complement(5971..6301)
7151..7207
/rpt_family="AluYb"
7667..7778
/rpt_family="CT-rich"
9309..9426
/rpt_family="MIR3"
complement(9716..10072)
/rpt_family="L2"
complement(11010..11072)
/rpt_family="MER103"
complement(11082..11330)
/rpt_family="MIR"
11594..11720
/rpt_family="MIR"
12048..12215
/rpt_family="MIR"
complement(12220..12679)
/rpt_family="MLTIC"
13429..13540
/rpt_family="L1MC/D"
13667..13777
/rpt_family="MER58A"
13778..13852
/rpt_family="MER58"
13932..13962
/rpt_family="AT-rich"
13966..15547
/rpt_family="L1ME1"
15708..16021
/rpt_family="AluSx"
16164..16294
complement(16491..16704)
/rpt_family="L1MC4a"

```

repeat_region complement(16710, .17026)
repeat_region /rpt_family="AluJo"
repeat_region complement(17031, .17250)
repeat_region /rpt_family="L1MC4a"
repeat_region complement(17309, .17492)
repeat_region /rpt_family="MER91B"
repeat_region 18089, .18404
repeat_region /rpt_family="AluJo"
repeat_region 18559, .18709
repeat_region /rpt_family="L1MB4"
repeat_region 18710, .18759
repeat_region /rpt_family="AT_rich"
repeat_region complement(18784, .19028)
repeat_region /rpt_family="MIR"
repeat_region 19480, .19612
repeat_region /rpt_family="MIR"
repeat_region complement(19648, .19723)
repeat_region /rpt_family="L2"
repeat_region 20222, .20401
repeat_region /rpt_family="MIR"
repeat_region 20431, .20738
repeat_region /rpt_family="AluSx"
repeat_region complement(20797, .20979)
repeat_region /rpt_family="MIR"
repeat_region complement(21159, .21650)
repeat_region /rpt_family="L1MB7"
repeat_region complement(22225, .22279)

```

Query Match 10.6%; Score 51.2; DB 8; Length 196859;
 Best Local Similarity 45.3%; Pred. No. 1.1;
 Matches 185; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

```

43 TATGAAATATTTTCTGATCTGATCTTACAGTTGATATGCTGCTGTAATTTT 102
145927 TATTTATATTTTATATATATTTTATATATTTTATATATTTTATATAT 145986
103 ATTGAGAGGATTTTAAACACACCTTAAAGCTTAAATTAATTAATTTCTCAT 162
145987 ATATATATATATATATATATATATATATATATATATATATATATATATA 146046
163 CTTTAAAGCACATTTAGCTGCTGAAGCAATTAACGCTGATATCTGTAACATC 222
146047 ATATATATATATATATATATATATATATATATATATATATATATATATA 146106
223 TCGCACTAAATCTTCTTACACGCGTCTGCTCTTCCAGGAGCTCCGAATATGCCA 282
146107 TTATATATATATATATATATATATATATATATATATATATATATATATAT 146166
283 CTATTTATCTGTGCAATTTTCCATTTATTTCCCTATTTGGTATTTGATGCGCTT 342
146167 ATATTTATATATATATATATATATATATATATATATATATATATATATAT 146226
343 AAATGACACCGATGATCTTCACTGTTGCGAGTTTGTCTTTGCTTCTTCTTAAAGCTC 402
146227 ATTTATATATTTTATATATATATATATATATATATATATATATATATATAT 146286
403 TTCAATTTATCTAAGCAAGTTTGTATATTTCAAAATATCTTGCTT 450
146287 ATATATTTATATATATATATATATATATATATATATATATATATATTT 146334

```

RESULT 14
 AC124696 193849 bp DNA linear HTG 29-Apr-2005
 LOCUS Mus musculus chromosome 16 clone RP24-173113, WORKING DRAFT
 DEFINITION
 AC124696
 ACCESSION AC124696.4 GI:62955023
 VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS Mus musculus (house mouse)
 SOURCE
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

REFERENCE

1 (bases 1 to 193849)
 Wilson, R.K.
 The sequence of Mus musculus clone
 Unpublished
 2 (bases 1 to 193849)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (15-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 193849)
 Wilson, R.K.
 Direct Submission
 Submitted (29-Apr-2005) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 On Apr 29, 2005 this sequence version replaced gi:24416033.

COMMENT

Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.edu
 Project Information
 Center project name: M_BB0173113

Summary Statistics

Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 191374 bases at least Q40
 Consensus quality: 192113 bases at least Q30
 Consensus quality: 192485 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1052: contig of 1052 bp in length
* 1053 1152: gap of unknown length
* 1153 2247: contig of 1095 bp in length
* 2248 2347: gap of unknown length
* 2348 3388: contig of 1041 bp in length
* 3389 3488: gap of unknown length
* 3489 4996: contig of 1508 bp in length
* 4997 5096: gap of unknown length
* 5097 6623: contig of 1527 bp in length
* 6624 6724: gap of unknown length
* 6724 9455: contig of 2731 bp in length
* 9455 9554: gap of unknown length
* 9555 80258: contig of 70704 bp in length
* 80259 80359: gap of unknown length
* 80359 193849: contig of 113491 bp in length.

```

FEATURES

source

/organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="16"
 /clone="RP24-173113"

misc_feature

/note="assembly_name:Contig23"
 1053..1152

gap

/estimated_length=unknown
 1153..2247
 /note="assembly_name:Contig30"
 2248..2347
 /estimated_length=unknown

```

misc_feature      2348..3388
                    /note="assembly_name:Contig42"
gap               3389..3488
                    /estimated_length=unknown
misc_feature      3489..4996
                    /note="assembly_name:Contig43"
gap               4997..5096
                    /estimated_length=unknown
misc_feature      5097..6623
                    /note="assembly_name:Contig45"
gap               6624..6723
                    /estimated_length=unknown
misc_feature      6724..9454
                    /note="assembly_name:Contig46"
gap               9455..9554
                    /estimated_length=unknown
misc_feature      9555..80258
                    /note="assembly_name:Contig47"
gap               80259..80358
                    /estimated_length=unknown
misc_feature      80359..193849
                    /note="assembly_name:Contig48"

ORIGIN
Query Match      10.4%; Score 50.4; DB 14; Length 193849;
Best Local Similarity 56.7%; Pred. No. 1.6;
Matches 93; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 43 TATGAATAATATTTTACTGTGATCTTACAAGTTGATATATGTTGTGTAACTTATTT 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84521 TATATATATATATATATATATACATACATACATATATATATATATATATATAT 84580

QY 103 ATTGAGAGCATATTTTAAACACCTTAGAAGCTTAAACCTTAAATTAATATTTCTTAT 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84581 ATTATATGATATATATATATATATATATATAGTAAACAACTTATATACATACATATATAT 84640

QY 163 CTTTAAAGGACATATATAGTGCGCTTAAAGCAATTAACAGTGATA 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84641 ATGTATGTACACATCGCAGATGAGATTTCACAAATTTGAGCTGACA 84684

RESULT 15
AC163463
LOCUS             244638 bp DNA linear HTG 01-JUL-2005
DEFINITION       Bos taurus clone CH240-125K20, *** SEQUENCING IN PROGRESS ***, 39
                    unordered pieces.
ACCESSION         AC163463
VERSION           AC163463.2 GI:68300099
KEYWORDS          HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE            Bos taurus (cow)
ORGANISM          Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Bovidae; Bovinae; Bos.
1 (bases 1 to 244638)
Muzny,D,Marie, Metzker,M,Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benhammed,F,
Biswal,K, Blair,J, Blankenhorn,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Drepper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,N, Forbes,L, Foster,M, Foster,P,
Fernandez,S, Finley,M, Flagg,N, Flieger,N, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Gebregiorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W,
Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,

```

```

TITLE             JOURNAL
REFERENCE         Woxley,K.C.
AUTHORS           Submitted (12-JUL-2005) Human Genome Sequencing Center, Department
TITLE            of Molecular and Human Genetics, Baylor College of Medicine, One
JOURNAL           Baylor Plaza, Houston, TX 77030, USA
                  3 (bases 1 to 244638)
REFERENCE         Cow Genome Sequencing Consortium.
AUTHORS           Submitted (01-JUL-2005) Human Genome Sequencing Center, Department
TITLE            of Molecular and Human Genetics, Baylor College of Medicine, One
JOURNAL           Baylor Plaza, Houston, TX 77030, USA
                  On Jun 29, 2005 this sequence version replaced gi:67514619.
                  The sequence in this assembly is a combination of BAC based reads
                  and whole genome shotgun sequencing reads assembled using Atlas
                  (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
                  in the feature table below represents a scaffold in the Atlas
                  assembly (a 'contig-scaffold'). Within each contig-scaffold,
                  individual sequence contigs are ordered and oriented, and separated
                  by sized gaps filled with Ns to the estimated size. The sequence
                  may extend beyond the ends of the clone and there may be sequence
                  contigs within a contig-scaffold that consist entirely of whole
                  genome shotgun sequence reads. Both end sequences and whole genome
                  shotgun sequence only contigs will be indicated in the feature
                  table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FHXA
Center clone name: CH240-125K20
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 229556 bases at least Q40
Consensus quality: 233265 bases at least Q30
Consensus quality: 236194 bases at least Q20
Estimated insert size: 239301; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

```

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 4622: contig of 4622 bp in length
* 4623 4672: gap of 50 bp
* 4673 8644: contig of 3972 bp in length
* 8645 8744: gap of unknown length
* 8745 12688: contig of 3944 bp in length
* 12689 12738: gap of 50 bp
* 12739 16581: contig of 3843 bp in length
* 16582 16631: gap of 50 bp
* 16632 22198: contig of 5567 bp in length
* 22199 22248: gap of 50 bp
* 22249 26048: contig of 3800 bp in length
* 26049 26098: gap of 50 bp
* 26099 37514: contig of 11416 bp in length
* 37515 38190: gap of 676 bp
* 38191 55334: contig of 17144 bp in length
* 55335 55384: gap of 50 bp
* 55385 75854: contig of 20470 bp in length
* 75855 76343: gap of 489 bp
* 76344 76344: contig of 1 bp in length
* 76345 76394: gap of 50 bp
* 76395 83875: contig of 7481 bp in length
* 83876 83925: gap of 50 bp
* 83926 94134: contig of 10209 bp in length
* 94135 94267: gap of 133 bp
* 94268 97884: contig of 3517 bp in length
* 97885 97885: gap of 50 bp
* 97886 108699: contig of 10865 bp in length
* 108700 109258: gap of 559 bp
* 109259 115882: contig of 6624 bp in length
* 115883 115932: gap of 50 bp
* 115933 121475: contig of 5543 bp in length
* 121476 121525: gap of 50 bp
* 121526 131673: contig of 10148 bp in length
* 131674 131723: gap of 50 bp
* 131724 140028: contig of 8305 bp in length
* 140029 140128: gap of unknown length
* 140129 151032: contig of 10904 bp in length
* 151033 151082: gap of 50 bp
* 151083 181837: contig of 30755 bp in length
* 181838 181887: gap of 50 bp
* 181888 184514: contig of 2627 bp in length
* 184515 184564: gap of 50 bp
* 184565 189352: contig of 4788 bp in length
* 189353 189402: gap of 50 bp
* 189403 195600: contig of 6198 bp in length
* 195601 195650: gap of 50 bp
* 195651 201941: contig of 6291 bp in length
* 201942 207379: contig of 5388 bp in length
* 207380 207429: gap of 50 bp
* 207430 209271: contig of 1842 bp in length
* 209272 209371: gap of unknown length
* 209372 211084: contig of 1713 bp in length
* 211085 211134: gap of 50 bp
* 211135 230592: contig of 9458 bp in length
* 230593 230642: gap of 50 bp
* 230643 233227: contig of 2585 bp in length
* 233228 233277: gap of unknown length
* 233278 234554: contig of 1127 bp in length
* 234555 236141: contig of 1587 bp in length
* 236142 226241: gap of unknown length

```

```

* 226242 227620: contig of 1379 bp in length
* 227621 227720: gap of unknown length
* 227721 229657: contig of 1937 bp in length
* 229658 229757: gap of unknown length
* 229758 231422: contig of 1665 bp in length
* 231423 231522: gap of unknown length
* 231523 233019: contig of 1487 bp in length
* 233020 233119: gap of unknown length
* 233120 235005: contig of 1886 bp in length
* 235006 235105: gap of unknown length
* 235106 236775: contig of 1570 bp in length
* 236776 240371: gap of unknown length
* 240372 240471: gap of 3596 bp in length
* 240472 244638: contig of 4167 bp in length.
* Location/Qualifiers
* 1. 244638
* /organism="Bos taurus"
* /mol_type="genomic DNA"
* /db_xref="taxon:9913"
* /clone="CH240-125K20"
* 4623. 4672
* /estimated_length=50
* gap 8645. 8744

```

Query Match 10.3%; Score 50; DB 14; Length 244638;
 Best Local Similarity 36.9%; Pred. No. 1.8;
 Matches 152; Conservative 0; Mismatches 260; Indels 0; Gaps 0;

```

QY 39 AAGATGATGAATTTTCTGCTGATCTTACAGCTGATATATGCTGTGCTACTT 98
DB 76496 AATTTTAAATTTTATTTATTTAAATTTNNAATTTTAAATTTTATTTATTTATTT 76555
QY 99 ATTATTTGAGAGTATTTTAAACACCTTAGAATTTAAATTTAAATTTATTTCT 158
DB 76556 ATATTTTAAAAAATTTTATTTATTTAATTTAATTTAATTTAATTTAATTTAATTT 76615
QY 159 CTATCTTAAAGGACATTTACGTGGCTAAGGCAATTTACGCTGATATTTAACT 218
DB 76616 AATATTTTAAATTTAATTTAATTTTATTTATTTAATTTAATTTAATTTAATTTA 76675
QY 219 CATGTCGACATTAATTTCTTCAACGCGTCTGCTTTCCAAAGGACCTCGAATAT 278
DB 76676 ATACATTTAAATTTAATTTAATTTTNNNTAATTTNNNTAATTTNNNTAATTTATTT 76735
QY 279 GCACTATTTATCTGCGCATTTCCAAATTTATATTCCTATTTGGGATTTGATGGCC 338
DB 76736 TTTTNNATTTNNNTTTTNTTTTNNNTTTNTTTTNTTTTNTTTTNTTTNTTTNTTT 76795
QY 339 GTTAAATGACGAGATTTGAATCTTCACTTGTGAGTTTGTCTTTGCTTCTTAA 398
DB 76796 NTTTNTTTTNTTNNNTTTTNTTNTTNNNTTNNCTTTTNTTNNNTTNTTNTTNTTNT 76855
QY 399 GGTCTTCATTTATCTAAAGCAAGTTTGTATTAATTTCAAAATCTTGTCTT 450
DB 76856 TTTTNTTTTNTTNNNTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 76907

```

Search completed: December 28, 2005, 10:11:47
 Job time : 2763 secs

this Page Blank (uspto)

XX Sequence 485 BP; 137 A; 90 C; 82 G; 176 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 485; DB 12; Length 485;
Best Local Similarity 100.0%; Pred. No. 3.3e-104;
Matches 485; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGGAGAGAGAGAAAGTGAATCTTCTGTTGCGTGAACAGATGAAAAATTTTACT 60
DB 1 GCGCGGAGAGAGAGAAAGTGAATCTTCTGTTGCGTGAACAGATGAAAAATTTTACT 60
QY 61 GGGATTAAGTAAAGTGAATGATGTTGGTGTGAATTTTATTGAGAGATTTTAA 120
DB 61 GGGATTAAGTAAAGTGAATGATGTTGGTGTGAATTTTATTGAGAGATTTTAA 120
QY 121 CACACCTTAAGTAAACTTAATAATAATTTCTATCTTTAAAGCACAATTA 180
DB 121 CACACCTTAAGTAAACTTAATAATAATTTCTATCTTTAAAGCACAATTA 180
QY 181 CGTGGCTAAGCAATTAGAGCTGATATAGTAAATCTCATGTGCGCAATAATTTCT 240
DB 181 CGTGGCTAAGCAATTAGAGCTGATATAGTAAATCTCATGTGCGCAATAATTTCT 240
QY 241 AACAGCGCTTCTCTCTTCCAGGAGCTCCGAATATGCCATTTATCTGTGCAAT 300
DB 241 AACAGCGCTTCTCTCTTCCAGGAGCTCCGAATATGCCATTTATCTGTGCAAT 300
QY 301 TCCAAATTAATTCCTCTATTTGGATTTGATGTGCGCTTAAATAGTACCGATTGA 360
DB 301 TCCAAATTAATTCCTCTATTTGGATTTGATGTGCGCTTAAATAGTACCGATTGA 360
QY 361 TCTTCACTTGTTCGAGTTTGTCTTTTGTCTCTAAGGCTTCAATTTATCTAAAGCA 420
DB 361 TCTTCACTTGTTCGAGTTTGTCTTTTGTCTCTAAGGCTTCAATTTATCTAAAGCA 420
QY 421 AGTTTGTATTAATCAAAATTAATCTTCTTCCATGATGAACTGCACTCAATGTGAG 480
DB 421 AGTTTGTATTAATCAAAATTAATCTTCTTCCATGATGAACTGCACTCAATGTGAG 480
QY 481 GTACC 485
DB 481 GTACC 485
RESULT 2
AD039660 standard; DNA; 4046 BP.
XX
AC AD039660;
XX
DT 29-JUL-2004 (first entry)
XX
DE Yeast chitin synthase 3 CAL1 gene.
XX
KM Heterologous gene; expression cassette; gene expression; Yeast; CAL1;
KM chitin synthase 3; gene; ds.
XX
OS Saccharomyces cerevisiae.
XX
PN US2004092020-A1.
XX
PD 13-MAY-2004.
XX
PF 20-JUN-2003; 2003US-00600230.
XX
PR 20-JUN-2002; 2002US-0390529P.
XX
PA (WILK/) WILKINSON J Q.
PA (MCBR/) MCBRIDE K.
PA (BERT/) BERTAIN S.
XX
PI Wilkinson JQ, Mcbride K, Bertain S;
XX

DR WPI; 2004-374960/35.
DR GENBANK; X57300.
XX
PT New recombinant expression cassette comprising a promoter that is
PT functional in plants, operably linked with a coding sequence and a non-
PT plant 3' termination sequence, useful for gene expression in plant cells.
XX
PS Disclosure; SEQ ID NO 81; 74dp; English.
XX
CC The present invention relates to heterologous genes comprising non-plant
CC 3'-termination sequences and plant expression cassettes incorporating the
CC heterologous genes. The invention is useful for gene expression in plant
CC cells. The present sequence is yeast chitin synthase 3 CAL1 gene. This
CC sequence is used in the invention.
XX
SQ Sequence 4046 BP; 1211 A; 739 C; 819 G; 1277 T; 0 U; 0 Other;
Query Match 97.9%; Score 474.8; DB 12; Length 4046;
Best Local Similarity 99.6%; Pred. No. 1.2e-101;
Matches 476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 6 CGAAGAGAGAAAGTGAATCTTCTGTTGCGTGAACAGATGAAAAATTTTACTGTGAT 65
DB 3546 CGAAGAGAGAAAGTGAATCTTCTGTTGCGTGAACAGATGAAAAATTTTACTGTGAT 3605
QY 66 ACTTACAAGTGAATATATGTTGTGTGTAATTTATTTTGAAGATTTTAAACACAC 125
DB 3606 ACTTACAAGTGAATATATGTTGTGTGTAATTTATTTTGAAGATTTTAAACACAC 3665
QY 126 CTTAACCTAAACCTTAATAATAATTTCTATCTTTAAAGGCATATTACGTGG 185
DB 3666 CTTAACCTAAACCTTAATAATAATTTCTATCTTTAAAGGCATATTACGTGG 3725
QY 186 CTAAAGCAATTAACAGCTGATATAGTAAATCTCATGTGCGCACTAAATTTCTTAACAC 245
DB 3726 CTAAAGCAATTAACAGCTGATATAGTAAATCTCATGTGCGCACTAAATTTCTTAACAC 3785
QY 246 GCGTTCGTCTCTTCCAAAGGAGCTCCGAATATGCCATTTATCTGTGCAATTTCCAA 305
DB 3786 GCGTTCGTCTCTTCCAAAGGAGCTCCGAATATGCCATTTATCTGTGCAATTTCCAA 3845
QY 306 TTTATATCCCTATTTGGATTTGATGTGCGCTTAAATAGTACCGATTGAATCTTC 365
DB 3846 TTTATATCCCTATTTGGATTTGATGTGCGCTTAAATAGTACCGATTGAATCTTC 3905
QY 366 ACTTGTGAGATTTGTCTTTTGTCTCTAAGGCTTCAATTTATCTAAAGCAAGTTT 425
DB 3906 ACTTGTGAGATTTGTCTTTTGTCTCTAAGGCTTCAATTTATCTAAAGCAAGTTT 3965
QY 426 TGTATTAATTAATAATTAATCTTCTTTCATGATGATTAACCTCAATGATGAGCTA 483
DB 3966 TGTATTAATTAATAATAATTAATCTTCTTTCATGATGATTAACCTCAATGATGAGCTA 4023
RESULT 3
ABL32718 standard; DNA; 5739 BP.
XX
AC ABL32718;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 691.
XX
KM Human; immune system disease; cytosine methylation; antiasthmatic;
KM antiarteriosclerotic; antianaemic; cytosaratic; nocrotic;
KM neutroprotective; anti-HIV; anticonvulsant; ophthalmological;
KM antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KM antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KM ds.
XX

OS Homo sapiens.
 XX
 XX WO200200928-A2.
 XX
 XX
 PD 03-JAN-2002.
 PF 02-JUL-2001; 2001WO-EP007537.
 XX
 XX 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 XX MPI; 2002-130909/17.
 DR
 XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 XX
 PS Claim 1; SEQ ID NO 691; 32pp + Sequence listing; German.
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 CC
 SQ Sequence 5739 BP; 1457 A; 38 C; 1217 G; 3027 T; 0 U; 0 Other;
 Query Match 9.5%; Score 46.2; DB 6; Length 5739;
 Best Local Similarity 45.9%; Pred. No. 0.5; Mismatches 228; Indels 1; Gaps 1;
 Matches 194; Conservative 0; Mismatches 228; Indels 1; Gaps 1;
 QY 31 TTGCGTACAGATGATGAAATATTTTCTGATGATCTTCAAGTTGATATGATG 90
 DB 729 TTGATTAATATTTTATTTTATTTTATTTTCTGTTTATTAAGTTCTTAAATTTTGT 788
 QY 91 TGTAACTATTTATTTGAGAGTATTTTAAACACCTTGAAGACTAAATTAATTAATA 150
 DB 789 TTTTAGTTTATGTTTGAATTTGTTAGGAAAGGTTTAAAT-AGTTTGTGATGTT 847
 QY 151 ATATTTCTATCTTTTAAAGGACATATTTACGTGGCTTAAGCAATTACACTGATTA 210
 DB 848 ATATTTTAAATTAATTTTGTGCAATTTTATTAATTAATTAATTTGTTATGTTGA 907
 QY 211 GTAAACCTCAGTGGCAGCTAAATTTCTTAAACAGCGTTCTGCTCTTCAAGGACT 270
 DB 908 TTAATTTTGTGCAAGATTTGTTTATTAATTTTGTGTTTGTGTTTATTTTATGTTT 967
 QY 271 CCGAATATGCACTATTTATCTGTGCAATTTCCAAATTAATTTCCCTATTTGGTATTT 330
 DB 968 TTTTATTAATTAATTTTGTGTTAGTATTTTGTGTTAGTATTTTAAATTTGTTATG 1027
 QY 331 ATGTGGCCGTTTAAATGTCACCGATGAACTTCTCAGTTGTGCAAGTTTGTCTTTGCT 390
 DB 1028 TTAAGTTTGTGTTATTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 1087
 QY 391 TCTCTAAAGTCTCAATTTATCTTAAGCAAGTTTGTATTAATCAAAATCTTTGCTT 450
 DB 1088 TTTTATTTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1147
 QY 451 TCT 453
 DB 1148 TAT 1150

ID AAS61080 standard; DNA; 6092 BP.
 XX
 XX AAS61080;
 AC
 XX
 XX
 DT 29-JAN-2002 (first entry)
 XX
 XX
 DE Human gene regulation-associated gene oligonucleotide #35.
 XX
 XX Human; Gene regulation-associated gene; severe combined immunodeficiency;
 XX cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 XX asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 XX renal disease; Preecclampsia; cardiac allograft vascular disease;
 XX colorectal cancer; thyroid cancer; oesophageal cancer; des; tumour;
 XX immunostimulant; cardiac; antiinflammatory; coagulant; antisthmatic;
 XX nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 OS
 XX Homo sapiens.
 XX
 XX WO200177375-A2.
 PN
 XX
 PD 18-OCT-2001.
 PF 06-APR-2001; 2001WO-EP003968.
 XX
 XX
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 XX (EPIC-) EPIGENOMICS AG.
 PA
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 XX MPI; 2002-017470/02.
 DR
 XX
 XX
 PT New nucleic acid sequences from chemically modified genes associated with
 PT gene regulation, useful for analyzing cytosine methylations for diagnosis
 PT and therapy of diseases e.g. severe combined immunodeficiency disease.
 XX
 PS Claim 1; SEQ ID NO 36; 26pp; English.
 CC The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The chemical
 CC pretreatment converts cytosine bases unmethylated at the 5-position to
 CC uracil or another base with hybridisation behaviour dissimilar to
 CC cytosine, to enable analysis of cytosine methylations. The DNA sequences,
 CC oligomers (or sets/arrays) and method are useful in the diagnosis of
 CC diseases (or predisposition to diseases) associated with gene regulation
 CC and in therapy of such diseases, by enabling analysis of the cytosine
 CC methylation patterns of such genes. Kits are provided. They are
 CC especially useful in diagnosis and therapy of e.g. severe combined
 CC immunodeficiency disease, cardiac disorders, haemophilia, solid tumours
 CC and cancer, Werner syndrome, asthma, HDR syndrome, Saethre-Chotzen
 CC syndrome, renal disease, preecclampsia, graft versus-host disease. The
 CC present sequence is a sequence included in the sequence data for this
 CC specification and is associated with the human gene regulation-associated
 CC genes. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WHO at ftp.wipo.int/pub/published_pat_sequences
 CC
 SQ Sequence 6092 BP; 1931 A; 122 C; 1140 G; 2899 T; 0 U; 0 Other;
 Query Match 9.5%; Score 46.2; DB 6; Length 6092;
 Best Local Similarity 46.8%; Pred. No. 0.5; Mismatches 203; Indels 3; Gaps 1;
 Matches 181; Conservative 0; Mismatches 203; Indels 3; Gaps 1;
 QY 62 TGAATCTTACAGTTGATATATGTTGTTGATCTTATTTATTTGAGAGTATTTAAC 121
 DB 4944 TGTATTTTAAAGTGTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 5003
 QY 122 ACACCTTAGACATTAATTAATTAATTTTCTATCTTTTAAAGCAGATTAAC 181

Db 5004 ATTTGTTTAATTAATTTGAAAATTTAGTTTTTATTATTAAAGTTTGAAATAT 5063
 QY 182 GTGGCTAAGGCAATTACAGCTGATATACGTAAAACTCATGTGCGACATAATCTCTTA 241
 Db 5064 AAAATGATTAATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 5123
 QY 242 ACAAGGCTTCTGCTCTTCTTCCAAAGGACCTCCGAATATGCCACTATTATATCTGGCATT 301
 Db 5124 TGTGTGATTTTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 5180
 QY 302 CCAATTTATATTTCCCTATTTGGGATTTGATGAGCGCTTAAATAGTACCGAATTAAT 361
 Db 5181 ATAGTTTAACTTTAAAGGTTATATAGATTAATTTATTTATTTATTTATTTATTTATTA 5240
 QY 362 CTTCACCTTGTCGAGTTTGTCTTTGCTCTCTTAAAGGCTCTCAATTTATCTAAAGCAA 421
 Db 5241 TTAGTATTAATTTATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTAA 5300
 QY 422 GTTTGTATTAATTCAAAATTAATCTTGGCT 448
 Db 5301 AGAATTAATTAATTAATTAATTAATTAATTTGAT 5327

RESULT 5

AB066998 standard; DNA; 37515 BP.

AB066998;

28-AUG-2002 (first entry)

Human angiogenesis associated polynucleotide SEQ ID NO 28.

Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
 inflammation; rheumatoid arthritis; diabetic retinopathy; antileucers;
 macular degeneration; inflammatory bowel disease; Crohn's disease;
 anti-neumatic; anti-arthritic; antidiabetic; antipsoriatic;
 antiarteriosclerotic; ds.

Homo sapiens.

MO20024654-A2.

13-JUN-2002.

06-DEC-2001; 2001WO-EP014320.

06-DEC-2000; 2000DE-01061338.

(EPIC-) EPIDENOMICS AG.

Schacht O;

WPI; 2002-500450/53.

New nucleic acid fragments from chemically treated angiogenesis-
 associated genes, useful for determining methylation status, e.g. in
 diagnosis or treatment of cancer.

Claim 1; SEQ ID NO 28; 41pp + Sequence Listing; German.

The invention relates to a nucleic acid (I) comprising a segment of 18
 bases of chemically pretreated DNA of angiogenesis-associated genes (II)
 having sequences (AB066997-AB067178) or their complements. (II), also
 related oligomers, are used to evaluate the methylation status and/or
 single-nucleotide polymorphisms, in angiogenesis-related genes, for
 diagnosis and treatment of eye diseases, proliferative retinopathy,
 neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
 diabetic retinopathy, macular degeneration caused by neovascularisation,
 psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
 Crohn's disease. Note: The sequence data for this patent did not form
 part of the printed specification, but was obtained in electronic format
 directly from WIPO at ftp.wipo.int/pub/published_pcc_sequences

SX Sequence 37515 BP; 11428 A; 425 C; 7165 G; 18497 T; 0 U; 0 Other;
 SQ Query Match 9.3%; Score 45; DB 6; Length 37515;
 Best Local Similarity 53.0%; Pred. No. 1.3;
 Matches 96; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 43 TATGAAATATTTTACTGTAATCTTACAACTGATATATGTTGTGTAATTT 102
 Db 14019 TATATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 14078
 QY 103 ATTTGAGGATTTTAAACACCTTGAACCTTAAATTAATTAATTTCTAT 162
 Db 14079 GTTGTGATTAATTAATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTT 14138
 QY 163 CTTTAAAGCACATTAATTAATTTGCTGCTAAGGCAATTAAGCTATATCTGTAATTAATG 222
 Db 14139 AATTTTGTATTTTTCGTAAGTTTAAAGAAATTAAGAGTATGATATTAATG 14198
 QY 223 T 223
 Db 14199 T 14199

RESULT 6

ABL32336 standard; DNA; 7201 BP.

ABL32336;

26-MAR-2002 (first entry)

Human immune system associated gene SEQ ID NO: 309.

Human; immune system disease; cytosine methylation; antiasthmatic;
 antiarteriosclerotic; anti-neumatic; cytosine methylation; nootropic;
 neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 anti-inflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 ds.

Homo sapiens.

WO200200928-A2.

03-JAN-2002.

02-JUL-2001; 2001WO-EP007537.

30-JUN-2000; 2000DE-01032529.

01-SEP-2000; 2000DE-01043826.

(EPIC-) EPIDENOMICS AG.

Olek A, Piepenbrock C, Berlin K;

WPI; 2002-130909/17.

Nucleic acid comprising fragment of chemically modified gene, useful for
 diagnosis and treatment of diseases associated with abnormal cytosine
 methylation.

Claim 1; SEQ ID NO 309; 32pp + Sequence Listing; German.

The present invention provides a number of human immune system associated
 genes which are modified by the methylation of cytosines. The sequences
 can be used in the diagnosis and treatment of immune system disorders,
 including eye diseases such as retinopathy, neovascular glaucoma and
 macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

XX Determining the degree of cytosine methylation in genomic DNA, useful for
 PT diagnosis and prognosis, comprises selective hybridization of amplicons
 PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridized to two classes, each with at least one member,
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
 CC degree of hybridization to both classes is determined from the label on
 CC the amplicon. From the ratio of labels hybridized to the two classes of
 CC oligomers, the degree of methylation is calculated. The method is used:
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
 CC and of a wide range of diseases, e.g. cancer, disorders of the central
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
 CC particularly by detecting mutations or single nucleotide polymorphisms
 CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
 CC investigating cell differentiation. The method allows the methylation
 CC status of many C residues to be determined simultaneously. ABQ13410-
 CC ABQ5412 represent genomic DNA sequences used to illustrate the method
 CC for determining the degree of cytosine methylation described in the
 CC disclosure of the invention

XX Sequence 517 BP; 91 A; 32 C; 38 G; 356 T; 0 U; 0 Other;

XX Query Match 9.2%; Score 44.4; DB 6; Length 517;

XX Best Local Similarity 44.6%; Pred. No. 0.86; Mismatches 216; Indels 0; Gaps 0;

XX Matches 174; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

XX 44 ATGAAATATTTTACTGATCTTACCAAGTATATATGTTGGTAACTTATTA 103

XX 63 ATTATATTTTTCGTAATTTTATTTAGATATTTTATTTTATTTTATTTTCGTT 122

XX 104 TTTGAGAGTATTTTAAACACCTTAGAATAAATAATTAATTTCTATTC 163

XX 123 TTTTCGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTTC 182

XX 164 TTTAAGGACATATACGCGTAAAGGCAATTACGATATACGTAAACATCAT 223

XX 183 GATATATTTATTTATTTTCGCGTGAAGTTATTCGTTTATTCGTTTTCGTT 242

XX 224 CGCCACTAAATCTCTAACAACGCTTCCTCTTTCCAAAGGACTCCGATATGCCAC 283

XX 243 ATGATATTTATTTTTCGTTATTTTCGTTATTTTCGTTATTTTATTTTTCGCG 302

XX 284 TATTTATCTGCGCATTTCCAAATTAATCCCTATGCGGATTTGATGCGCGTTTA 343

XX 303 GATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAA 362

XX 344 AATAGCAGCGATTTGATCTTCGCTGCTGAGTTTTCGTTTCGCTCAAGGCTC 403

XX 363 TATTTATATTTTTCGTTTATTTTTCGTTTATTTTTCGTTTATTTTATTTTAA 422

XX 404 TCAATTTATCTAAGCAAGTTTGTATTAAT 433

XX 423 CGTTTTCGTTTATTTTCGTTTATTTTCGTTTAT 452

XX RESULT 9

XX ACL35363/c

XX ACL35363 standard; cDNA; 2000 BP.

XX ACL35363;

XX 02-JUN-2005 (first entry)

XX R1ce stress-regulated promoter SEQ ID NO:13926.

KM ss; abiotic stress tolerance; transgenic plant; plant; cereal;

XX agriculture.

XX Oryza sativa.

XX WO2003008540-A2.

XX 30-JAN-2003.

XX 21-JUN-2002; 2002WO-US019668.

XX 22-JUN-2001; 2001US-0300112P.

XX 24-AUG-2001; 2001US-0314662P.

XX 26-SEP-2001; 2001US-0325277P.

XX 21-NOV-2001; 2001US-0332132P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Krepes J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F;

XX Moughamer T, Provart N, Rieke D, Zhu T;

XX WPI, 2003-248011/24.

XX Claim 48; SEQ ID NO 13926; 89pp; English.

XX The invention relates to novel abiotic stress responsive polynucleotides

XX and polypeptides. Also disclosed are vectors, expression cassettes, host

XX cells, and plants containing such polynucleotides. Also disclosed are

XX methods for using the polynucleotides and polypeptides to alter the

XX responsiveness of a plant to abiotic stress. The invention is useful in

XX agriculture. The nucleic acid is useful for determining whether a test

XX plant has been exposed to an abiotic stress condition. It is also useful

XX for selecting an agent that alters abiotic stress regulated

XX polynucleotide expression in a plant cell, and to identify a homolog or

XX ortholog to an abiotic stress responsive polynucleotide. The nucleic acid

XX molecule and the polypeptide encoded by it are useful in altering the

XX responsiveness of a plant to an abiotic stress, such as cold stress, salt

XX stress, osmotic stress or any of their combinations. The present sequence

XX is used in the exemplification of the invention

XX Sequence 2000 BP; 540 A; 344 C; 416 G; 428 T; 0 U; 272 Other;

XX Query Match 9.0%; Score 43.6; DB 11; Length 2000;

XX Best Local Similarity 12.8%; Pred. No. 1.7; Mismatches 147; Indels 0; Gaps 0;

XX Matches 43; Conservative 147; Mismatches 146; Indels 0; Gaps 0;

XX 9 AAGAGGAAAGTACTCTTCGTTGCGTAGACAGATGAAATATTTTACTGTATCT 68

XX 373 AAGGAAGTCGTTTAAAGGCAAGTATTTTAAAGGCAATTTTAAACACCTT 314

XX 69 TCAAGTATGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 128

XX 313 YTSMAAGKMWMSXTTARNSWAMRWTKMTRAAISKWKKMMWTRAAATTTM 254

XX 129 AAGCAATTAAGCTATTAATAATATTTCTCATCTTTTAAAGGCAATTTTAAAGGCTA 188

XX 253 TMMMTMMWARKKTRTMMKTYAMWTKYAAAMAATATWTGATGAMWTRSMAGMAA 194

XX 189 AAGCAATTAAGCTATTAATAATATTTCTCATCTTTTAAAGGCAATTTTAAAGGCTA 248

XX 193 AKWYTKWTRBAAWMTYRTMTKMTSWCYTMMWRMMWWTWTRWRRKMRARGRA 134

XX 249 TTCTGCTCTTTCGAAGGACTCCGAATATGCAATATTTATCTGATTTTCAATTT 308

XX 133 GTMMWRRKAYMSWTYTKKKKMYMRYKRYCKPRAMMSWCYTKSMWYKMSWRSR 74

XX 309 ATATTCCTCATTTGATTTGATGTCGCGCTTTAA 344

XX 73 MSGSAGYMTKYAKYMMWMTTRMTKTRTKTSWAMMM 38

```

RESULT 10
ACL37108/c
ID ACL37108 standard; cDNA; 2000 BP.
XX
AC ACL37108;
XX
DT 02-JUN-2005 (first entry)
XX
DE Rice stress-regulated promoter SEQ ID NO:15671.
XX
KM ss; abiotic stress tolerance; transgenic plant; plant; cereal;
XX agriculture.
XX
OS Oryza sativa.
XX
PN MO2003008540-A2.
XX
PD 30-JUN-2003.
XX
PF 21-JUN-2002; 2002MO-US019668.
XX
PR 22-JUN-2001; 2001US-0300112P.
XX
PR 24-AUG-2001; 2001US-0314662P.
XX
PR 26-SEP-2001; 2001US-0325277P.
XX
PR 21-NOV-2001; 2001US-0332132P.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Krepe J, Briggs SP, Cooper B, Glazebrook J, Goff SA, Katagiri F,
PI Moughamer T, Frovaart N, Rieke D, Zhu T;
XX
DR WPI; 2003-248011/24.
XX
PT New stress-responsive nucleic acid, useful for altering the
PT responsiveness of a plant, e.g. cereal, to an abiotic stress such as cold
PT stress, salt stress or osmotic stress.
XX
PS Claim 48; SEQ ID NO 15671; 89bp; English.
XX
CC The invention relates to novel abiotic stress responsive polynucleotides
CC and polypeptides. Also disclosed are vectors, expression cassettes, host
CC cells, and plants containing such polynucleotides. Also disclosed are
CC methods for using the polynucleotides and polypeptides to alter the
CC responsiveness of a plant to abiotic stress. The invention is useful in
CC agriculture. The nucleic acid is useful for determining whether a test
CC plant has been exposed to an abiotic stress condition. It is also useful
CC for selecting an agent that alters abiotic stress regulated
CC polynucleotide expression in a plant cell, and to identify a homolog or
CC ortholog to an abiotic stress responsive polynucleotide. The nucleic acid
CC molecule and the polypeptide encoded by it are useful in altering the
CC responsiveness of a plant to an abiotic stress, such as cold stress, salt
CC stress, osmotic stress or any of their combinations. The present sequence
CC is used in the exemplification of the invention
XX
SQ Sequence 2000 BP; 510 A; 350 C; 268 G; 518 T; 0 U; 354 Other;

```

```

Query Match 8.9%; Score 43; DB 11; Length 2000;
Best local similarity 12.2%; Pred. No. 2.3; Indels 4; Gaps 2;
Matches 55; Conservative 206; Mismatches 186;
QY 9 AAGGAGAAAGTACTCTTCGTTGCGTAGACAGTATGAAATATTTTACTGTGATACT 68
DB 451 ARKAWGKARWYCKRTTAAVATMTTTRRRRAAMWMTGKRKRBSMTMRMSRTWTW 392
QY 69 TACAATGATATATNGTGTGTGTAACCTTA--TTATTTGAGAGGATATTTAAACACC 126
DB 391 YKYYTWTTRRYSRRRTGKAKKYMTWMMWGCMSYTKARWMAAYKSWMAVYAWT 332
QY 127 TTGAGACTAAACTATATATATATATATTTCTATCTTAAAGGACATTAAGCTGGC 186
DB 331 KGRTRYYMKRAAYVTRMTATRSATWCTGWTYTSKATTTKMRMGVYAKATTTMKRKY 272

```

```

QY 187 TAAGCAATTACAGCTGATATATCTGTAAAACTCATGTCGCCACTAAATTTCTTACACG 246
DB 271 YAGGRWWWYKKWASRTWKMCMYKMYMMATKKSSWMSMTTAMKSMWYKKWTYYTMTS 212
QY 247 CGTTCGTCTCTTCCAGGAGGACTCCGAATNCCACT--ATTATGCTGGCATTTCCA 304
DB 211 TYWTWMMWRCPAGWMMWGGAAKRAAATKMTRTYRKMKMMRYRYKMSWAGMAMRKR 152
QY 305 ATTATATTTCCCTTATTTGGTATTTGANGTGGCCGCTTAAATAGTCACGATTTGATCTT 364
DB 151 YTTWMMWMMYRCAMWMSAYWRSWRMTTXXSKMYKMMWRRTYBAKAYSRMMWK 92
QY 365 CACTGTTCGAGTTTGTCTTCTTCTCTCTTAAAGGCTTCAATTATTTAAACAAGTT 424
DB 91 KCRTRMKGTWRTKTKWMMWRWYKKRTYRKMSAKMWTWRSAGAYRRBWRMAYSRBSW 32
QY 425 TTGTATAATTCAAAATTAATCTTCTTTCTCC 455
DB 31 MKRTSYYYTKWKKWKMSSGYTWRYYTTTMS 1

```

```

RESULT 11
ADZ71063/c
ID ADZ71063 standard; DNA; 847 BP.
XX
AC ADZ71063;
XX
DT 14-JUL-2005 (first entry)
XX
DE Human chromosome 2 contig DNA SEQ ID NO 187.
XX
KM ds; matrix attachment region; MAR; protein production.
XX
OS Homo sapiens.
XX
PN WO2005040377-A2.
XX
PD 06-MAY-2005.
XX
PF 22-OCT-2004; 2004MO-BP011974.
XX
PR 24-OCT-2003; 2003US-0513574P.
XX
PR 06-FEB-2004; 2004EP-00002722.
XX
PA (SELEX-) SELEXIS SA.
XX
PI Mermoud N, Giroud PA, Bucher P, Nguyen D, Calabrese D, Sauvy D;
PI Puttini S;
XX
DR WPI; 2005-333507/34.
XX
PT New purified and isolated DNA sequence having protein production
PT increasing activity comprises a bent DNA element and a binding site for a
PT DNA binding protein, useful for increasing protein production activity in
PT eukaryotic host cell.
XX
PS Disclosure; SEQ ID NO 187; 282bp; English.
XX
CC The invention relates to a purified and isolated DNA sequence having
CC protein production increasing activity comprising at least one bent DNA
CC element, and at least one binding site for a DNA binding protein. The
CC purified and isolated DNA sequence comprising a first and a second
CC isolated matrix attachment region (MAR) nucleotide sequence, which is a
CC MAR nucleotide sequence selected from a purified and isolated DNA
CC sequence above, a purified and isolated MAR DNA above, a purified and
CC isolated clysmAR element and/or fragment, a synthetic MAR sequence, a
CC sequence complementary to it, its molecular chimera, or its combinations
CC and variants, is useful for increasing protein production activity in a
CC eukaryotic host cell. The present sequence represents a human chromosome
CC 2 contig DNA.
XX
SQ Sequence 847 BP; 501 A; 17 C; 10 G; 319 T; 0 U; 0 Other;

```

Query Match	8.7%	Score 42.4	DB 14	Length 847
Best Local Similarity	45.6%	Pred. No. 2.8	Indels 1	Gaps 1
Matches 186	Conservative 0	Mismatches 221		
QY	43	TATGAAATATTTTCTGATCTGATCTTACCAAGTTGATATGTTGTGTGTAACCTTATTT	102	
DB	454	TATTTATATATATTTTATATATATATATATTTTATATATATATATATATATATATATATTTT	395	
QY	103	ATTGAGAGGTAATTTTAAACAACCTTAGACTPAAACCTTAATPAAATATTTCTCTAT	162	
DB	394	TATATATTTATATATATTTATATA-TATATTTAATATATTTTATATATATATATATATATCTAT	336	
QY	163	CTTTAAAGCACATATATTACGTGGCTTAAGCAATTTACAGCTGATATCTGAAACCTCATG	222	
DB	335	TATATATATTTTATATATATTTATATGATATATATATATATATATATATATATATCTTT	276	
QY	223	TGGCACATAATTTCTTCTPAAACAGCGCTCTGTCTTTCCAGAGGACTCCGAATATGCCA	282	
DB	275	ATATATTTTATGATATATATATTTTATACATATATATTTATGATATATATATCTATATAT	216	
QY	283	CTATTATCTGTGGCAATTTCCAAATTTATATTTCCCTATTTGGTATTTGATGTGGCGTTT	342	
DB	215	TATATGATATATATATTTTATACATATATATTTTATGATATATATATATATATATATAT	156	
QY	343	AAATAGTACCGATTTGAATCTTCACTGTTCGAGTTTGTCTTTGCTTCTPAAAGTTC	402	
QY	155	ATATATTTTATATATGATATATATTTTATATATTTTATATATTTTATATATATATTTT	96	
QY	403	TTCATTTTATCTAAGCAAGTTTGTATATTTCAATTTCAATTTCTGCTT	450	
DB	95	GTAATATATTTTATATATTTTATATATATTTTATATATATTTTATATAT	48	
RESULT 12				
ABV99761/c				
ID	ABV99761	standard; DNA; 22617 BP.		
XX	ABV99761;			
XX	24-FEB-2003	(first entry)		
XX	Human PFKEB2 gene.			
XX	Human; 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2; PFKEB2;			
KW	cytostatic; antidiabetic; gene therapy; cancer; diabetes; SNP; gene; ds;			
KM	single nucleotide polymorphism.			
XX	Homo sapiens.			
OS	Location/Qualifiers			
XX	Key	replace(1342,G)		
FH	variation	/*tag= a		
FT	variation	/standard name= "single nucleotide polymorphism"		
FT	variation	replace(1366,T)		
FT	variation	/*tag= b		
FT	variation	/standard name= "single nucleotide polymorphism"		
FT	variation	replace(1686,A)		
FT	variation	/*tag= c		
FT	variation	/standard name= "single nucleotide polymorphism"		
FT	CDS	1332..20748		
FT	CDS	/*tag= d		
FT	CDS	/product= "PFKEB2"		
FT	CDS	/note= "Contains 13 introns"		
FT	exon	3132..3216		
FT	exon	/*tag= e		
FT	exon	/number= 1		
FT	intron	3217..10307		
FT	intron	/*tag= f		
FT	intron	/number= 1		
FT	variation	replace(10169,A)		
FT	variation	/*tag= g		
FT	variation	/standard name= "single nucleotide polymorphism"		

FT	variation	replace(10214,C)
FT		/*tag= h
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	10308. .10433
FT		/*tag= i
FT		/number= 2
FT	intron	10434. .10975
FT		/*tag= j
FT		/number= 2
FT	variation	replace(10965,A)
FT		/*tag= k
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	10976. .11072
FT		/*tag= l
FT		/number= 3
FT	intron	11073. .11498
FT		/*tag= m
FT		/number= 3
FT	exon	11499. .11565
FT		/*tag= n
FT		/number= 4
FT	variation	replace(11545,A)
FT		/*tag= o
FT		/standard_name= "Single nucleotide polymorphism"
FT	intron	11566. .11702
FT		/*tag= p
FT		/number= 4
FT	exon	11703. .11777
FT		/*tag= q
FT		/number= 5
FT	intron	11778. .12128
FT		/*tag= r
FT		/number= 5
FT	exon	12129. .12185
FT		/*tag= s
FT		/number= 6
FT	intron	12186. .13405
FT		/*tag= t
FT		/number= 6
FT	exon	13406. .13530
FT		/*tag= u
FT		/number= 7
FT	intron	13531. .15872
FT		/*tag= v
FT		/number= 7
FT	exon	15873. .16080
FT		/*tag= w
FT		/number= 8
FT	intron	16081. .16538
FT		/*tag= x
FT		/number= 8
FT	exon	16539. .16685
FT		/*tag= y
FT		/number= 9
FT	intron	16686. .17798
FT		/*tag= z
FT		/number= 9
FT	exon	17799. .17903
FT		/*tag= aa
FT		/number= 10
FT	variation	replace(17939,A)
FT		/*tag= ab
FT		/standard_name= "Single nucleotide polymorphism"
FT	intron	17904. .18657
FT		/*tag= ad
FT		/number= 10
FT	variation	replace(17911,A)
FT		/*tag= ac
FT		/standard_name= "Single nucleotide polymorphism"
FT	exon	18658. .18787
FT		/*tag= ae
FT		/number= 11
FT	intron	18788. .19565

FT		/tag= af	/number= 11
FT	exon	19566.	.19628
FT		/tag= ag	/number= 12
FT	intron	19629.	.19885
FT		/tag= ah	/number= 12
FT	exon	19886.	.19950
FT		/tag= ai	/number= 13
FT	intron	19951.	.20580
FT		/tag= aj	/number= 13
FT	exon	20581.	.20748
FT		/tag= ak	/number= 14
FT	variation	replace(20901,T)	
FT		**tag= al	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(21098,G)	
FT		/tag= am	
FT	variation	/standard_name= "Single nucleotide polymorphism"	
FT		replace(21101,G)	
FT		**tag= an	
FT		/standard_name= "Single nucleotide polymorphism"	
XX			
PN	WO200194363-A2.		
PD	13-DEC-2001.		
XX			
PE	07-JUN-2001; 2001WO-US018458.		
XX			
PR	07-JUN-2000; 2000US-0209935P.		
XX	(GENA-) GENAISSANCE PHARM INC.		
PA	Duda A, Kazemi A, Koshy B;		
PI	WPI; 2002-566434/60.		
DR	P-PsDB; ABP57966.		
XX			
PS	New 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2 (PFKFB2) gene variants, for improving efficiency and reliability in the development of drugs for treating diseases associated with PFKFB2 activity e.g. cancer.		
XX			
PS	Example 1; Fig 1; 95bp; English.		
XX			
CC	The invention relates to a novel human 6-phosphofructo-2-kinase/ fructose		
CC	-2,6-biphosphataae 2 (PFKFB2) isogene. The PFKFB2 of the invention has		
CC	cystostic and antibalabac activity. The polynucleotides may have a use		
CC	in gene therapy. The identified candidate agents targeting PFKFB2, are		
CC	useful for treating cancer and diabetes. The methods of the invention are		
CC	useful for improving the efficiency and reliability of several steps in		
CC	the discovery and development of drugs for treating diseases associated		
CC	with PFKFB2 activity. The present sequence encodes the PFKFB2 of the		
CC	invention		
XX			
SQ	Sequence 22617 BP; 5587 A; 4922 C; 5366 G; 6741 T; 0 U; 1 Other:		
	Query Match	8.7%; Score 42.2; DB 6; Length 22617;	
	Best Local Similarity	57.0%; Pred. No. 5.5;	
	Matches	77; Conservative 0; Mismatches 58; Indels 0; Gaps 0;	
OY	43 TATGAAATATTTTACTGTGATTAACAAGTGATATATGCTTGCTGAACCTTATT	102	
DB	9535 TATATTAAGAATATATATTAATATATATAAAGATATATTAAGAATATATTAATAATATAT	9476	
OY	103 ATTTGAGAGGTATTTTAAACACCCTTAGAACCTTAACCTTAATTAATAATATTTCTCAT	162	
DB	9475 AATATTAAGATATTTATATATATCTTAAATATATATATTTAAAGATATATGATATATATNG	9416	
OY	163 CTTTAAAGGCACATA	177	

Db 9415 ATATATGATATATATA 9401

RESULT 13
ID AAS46704
AAS46704 standard; DNA; 14615 BP.
XX AAS46704;
XX
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #427.
DE
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
XX Homo sapiens.
OS
XX WO200168912-A2.
XX
XX 20-SEP-2001.
PD
XX 15-MAR-2001; 2001WO-EP002955.
PF
XX 15-MAR-2000; 2000DE-01013847.
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI WPI; 2001-602752/68.
XX
XX
XX Fragments of chemically modified genes associated with tumor suppressor
PT genes and oncogenes, useful in designing primers and probes for analyzing
PT diseases associated with cytosine methylation state e.g. cancer.
XX
XX
XX Claim 1; SEQ ID NO 427; 27bp; English.
PS
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and oncogenes
CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
CC 500 are missing from the sequence listing) sequences (Ss) and sequences
CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
CC probes for detecting the cytosine methylation state and/or single
CC nucleotide polymorphisms and also to be used in an array for analysing
CC diseases associated with CpG dinucleotides e.g. cancers and tumour. The
CC probes can also be used in a method for ascertaining genetic and/or
CC epigenetic parameters for the diagnosis and/or therapy of existing
CC diseases or the predisposition to specific diseases, by analysing
CC cytosine methylations. The parameters may be compared to another set of
CC genetic and/or epigenetic parameters, the differences serving as basis
CC for diagnosis and/or prognosis events which are disadvantageous to
CC patients. The present sequence is one of the 533 genomic sequences
CC derived from tumour suppressor genes and oncogenes. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 14615 BP; 4072 A; 253 C; 323 G; 7057 T; 0 U; 0 Other;
SQ

Query Match 8.7%; Score 42; DB 4; Length 14615;
Best Local Similarity 56.5%; Pred. No. 5.6;
Matches 78; Conservative 0; Mismatches 60; Indels 0; Gaps 0.

43 TATGAAAAATTTTACTGTCGATACCTTACAAAGTGTGATATATGCTTGCTGTGACTTATTT 102

oligomers, are useful for diagnosis of diseases associated with gene repair, specifically ataxia telangiectasia, aging, Bloom's syndrome, Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome, immunodeficiency, trichothiodystrophy, Fanconi's anaemia, solid tumours and cancer, particularly by determining status of cytosine methylation and/or by detecting single-nucleotide polymorphisms. Determination of individual methylation patterns may allow development of individualised therapies. The sequences given in records AB192192-AB192335 represent chemically pre-treated DNA fragments from genes associated with DNA repair, and their complements. Note: The sequence data for this patent is not represented in the specification, but is based on sequence information supplied by the European Patent Office

XX Sequence 6801 BP; 2041 A; 163 C; 1563 G; 3034 T; 0 U; 0 Other;

Query Match 8.6%; Score 41.6; DB 6; Length 6801;

Best Local Similarity 44.1%; Pred. No. 6.1; Mismatches 0; Gaps 0;

Matches 173; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

```
QY 60 TGTGATCTTACAAAGTTGATATATGCTGTGCTGTAACCTTATTTTGGAGAGGTAATTTA 119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5358 TTGTAAATTTAGATATTTAATAGGTTTGTATTTAGATTTATTAATAATATTTTAG 5417

QY 120 ACACACTTAGAAGCTAAACTTAATAATAATTTCTATCTTAAAGCACAATTT 179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5418 TTTAATTAATGAAAGATTAATTAATTAATAAGAGATTAATTAATTAATTAATTTT 5477

QY 180 ACGTGCTAAGCAATTACAGCTGATATCTGTAATACTGATGCGCACTAATTTCTTC 239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5478 TCGAAGAGATGATTTGATGATTAATAATAATAATAATTAATTTATTTAGTTATTA 5537

QY 240 TAAACAGCGTCTGCTCTTTTCCAGGAGCTCCGAATATGCCATTTATCTGTGGCAT 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5538 GGGAAATGTAATTAAGTATAGTATGATATTTATTTATTTTGTAGTAAATTTT 5597

QY 300 TTCCATTTATATTCCTTATTTGGGATTTGATGATGCGCTTAATAATAGCACCATTTGA 359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5598 TTGTAAATTTAATTTGTAATAATAATTTTATTTTATTTTATTTATTTATTTT 5657

QY 360 ATCTGACTTGTGAGTTTGTCTTTGCTTCTTAAGGCTTCAATTTATCTTAAGC 419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5658 TTTTATTTTATTTTTCGTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTT 5717

QY 420 AAGTTTGTAAATTCAAAATACCTTGTCTTT 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5718 TTTTATTTATGATATTCGTTTGTATTT 5749
```

Search completed: December 28, 2005, 09:24:31
Job time : 499 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 08:09:16 ; Search time 3840 Seconds
(without alignments)
5909.306 Million cell updates/sec

Title: US-10-600-230-1

Sequence: 1 gcgcgcgcgaaggagaagt.....cctccaatgatgaggtacc 485

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hlc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66.8	13.8	495	9 B2302050	B2302050 K01625.p1
2	55.6	11.5	1001	10 CNS0064G	AL062781 Drosophila
3	51.8	10.7	1101	10 CNS0039G	AL063921 Drosophila
4	48.8	10.1	1101	10 CNS006VL	AL069706 Drosophila
5	47.8	9.9	1191	10 CL104745	CL104745 ISB1-4311
6	47.6	9.8	705	9 BH945927	BH945927 obu88a02.
7	47	9.7	1101	10 CNS0106X	AL089595 Drosophila
8	46.8	9.6	1101	10 CNS016LT	AL106886 Drosophila
9	46.2	9.5	1101	10 CNS0182P	AL108811 Drosophila
10	46.2	9.5	1161	1 A0929051	A0929051 A0929051
11	45.8	9.4	1079	1 A0928310	A0928310 A0928310
12	45.4	9.4	1014	10 DU000686	DU000686 33624 To
13	45.2	9.3	698	8 CX173493	CX173493 A09_69-22
14	45.2	9.3	718	9 CE216145	CE216145 tigr-g88-
15	45	9.3	215	1 AV112242	AV112242 AV112242
16	44.5	9.3	1484	9 CC229324	CC229324 CH261-183
17	44.6	9.2	717	9 B2454798	B2454798 BOMAI15TR
18	44.6	9.2	891	10 C2977617	C2977617 191754 To
19	44.6	9.2	914	10 CNS002JY	AL097768 Drosophila
20	44.6	9.2	1101	10 CNS001PB	AL060732 Drosophila
21	44.4	9.2	974	10 CNS001TT	AL075432 Drosophila
22	44	9.1	1204	10 CNS016E2	AL106628 Drosophila

23	43.8	9.0	928	10 CNS00DXY	AL071865 Drosophila
24	43.6	9.0	549	6 CD441919	CD441919 EL01N0402
25	43.6	9.0	733	10 CE692306	CE692306 tigr-g88-
26	43.6	9.0	868	10 CNS010YV	AL099601 Drosophila
27	43.6	9.0	966	6 CD388381	CD388381 AGENCOURT
28	43.4	8.9	332	3 BI814443	BI814443 PEST0a2
29	43.4	8.9	921	10 AG552152	AG552152 Mus muscu
30	43.4	8.9	987	10 CNS014PQ	AL104456 Drosophila
31	43.4	8.9	1057	8 DN782797	DN782797 92300960
32	43.4	8.9	1101	10 CNS0039R	AL063932 Drosophila
33	43.2	8.9	547	3 BM535026	BM535026 EST588048
34	43.2	8.9	555	9 A0451200	A0451200 HS_5178_A
35	43.2	8.9	750	10 CNS011ID	AL100303 Drosophila
36	43.2	8.9	811	1 AL514901	AL514901 AL514901
37	43.2	8.9	1258	10 CL508886	CL508886 SAIL_804_
38	43	8.9	434	9 BH183920	BH183920 T7 end of
39	43	8.9	434	11 CNS070Z8	AL208070 T7 end of
40	43	8.9	879	10 CL486692	CL486692 SAIL_440_
41	43	8.9	1126	9 CC215864	CC215864 CH261-117
42	43	8.9	1203	10 CNS015WU	AL106608 Drosophila
43	42.8	8.8	415	9 A0203034	A0203034 RPEC11-63
44	42.8	8.8	422	3 BI815751	BI815751 PEST0a3
45	42.8	8.8	697	9 BH700251	BH700251 BOMEN12TF

ALIGNMENTS

RESULT 1	B2302050	495 bp	DNA	linear	GSS 31-OCT-2002
LOCUS	B2302050				
DEFINITION	K01625.p1 Kluveromyces delphensis Random Genomic Library				
ACCESSION	B2302050				
VERSION	B2302050.1	GI:24448520			
SOURCE	GSS.				
ORGANISM	Kluveromyces delphensis				
REFERENCE	1 (bases 1 to 495)				
AUTHORS	Wong,S., Fares,M.A., Zimmermann,W., Butler,G. and Wolfe,K.H.				
TITLE	Evidence from comparative genomics for a complete sexual cycle in the 'asexual' pathogenic yeast Candida glabrata				
JOURNAL	Genome Biol. 4 (2), R10 (2003)				
COMMENT	12620120				
CONTACT	Contact: Wong S				
DEPARTMENT	Department of Genetics, Smurfit Institute				
LOCATION	Trinity College Dublin				
ADDRESS	Dublin 2, Ireland				
TEL	Tel: 353 1 6082319				
FAX	Fax: 353 1 6798558				
EMAIL	Email: swong@tcd.ie				
CLASS	Class: plasmid ends.				
FEATURES	Location/Qualifiers				
SOURCE	1..495				
ORGANISM	/organism="Kluveromyces delphensis"				
MOL_TYPE	/mol_type="genomic DNA"				
SCRATCH	/scratch="CBS 2170"				
DB_XREF	/db_xref="taxon:51657"				
CLONE	/clone="K01625"				
LIBRARY	/clone_lib="Kluveromyces delphensis Random Genomic Library"				
ORIGIN					
Query Match	13.8%	Score 66.8;	DB 9;	Length 495;	
Best local similarity	56.5%;	Pred. No. 7.1e-05;			
Matches	166;	Conservative 0;	Mismatches 122;	Indels 6;	Gaps 2;
QY	195	TTACAGCTGATATACCTGTAACCTCATGTCGACCACTAATTCCTTCAACAGCGGCTCTG	254		
DB	82	TCAAAATTCATACAGTAACCTTACGACCCATTAATTTCTCCAGCAGCTCTTCTGT	141		

Qy	255	CTCTTTCCAAAGGAGCTCCGAATATGACCACTATTATCTGTGGCAATTCATTTATATATC	314
Db	142	TCGTTGCCAAGGACACCAAAAGTCCGCTATATATCTTTGGCATTTCAACAGGTATTC	201
Qy	315	CCCTATTTGGATTTTGTATGTAGTGGCCGCTTTAA-----ATATGACCGAATGAATCTTCACCTT	369
Db	202	GTTAAGGTACCTTCATTAATTTGTGATTAATATTCCTTATTCATTAATTCAGATTAGTAT	261
Qy	370	GTTGAGTTTGTCTTTTGGCTTCTCTTAAGGCTTCATTTA-TCATAAGCAAGTTTGT	428
Db	262	ACGACAGATATGAGGACTCTCAGTCAATTAATTCATTAAAGCAATTCAAAGATTTTGGGT	321
Qy	429	ATTAATTCAAAATACCTTTCCTTTCTCCATGACTTGAACTTCMAATGATGAGGT	482
Db	322	ATATTTCTAAATGCTGTGTGACAGATCCATGTTTCAACCGCAAAAGATGATGT	375
RESULT 2			
LOCUS	CNS0064G		
DEFINITION	CNS0064G	1001 bp	DNA
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC #		
DEFINITION	BACR13009 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
DEFINITION	AL062781		
DEFINITION	AL062781.1	GI:4943971	
DEFINITION	GSS.		
DEFINITION	Drosophila melanogaster (fruit fly)		
DEFINITION	Drosophila melanogaster		
DEFINITION	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
DEFINITION	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
DEFINITION	Ephydroidea; Drosophilidae; Drosophila.		
DEFINITION	1 (bases 1 to 1001)		
DEFINITION	Genoscope.		
DEFINITION	Direct Submission		
DEFINITION	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :		
DEFINITION	BP 191 91006 EVRY cedex - FRANCE (E-mail : seget@genoscope.cns.fr		
DEFINITION	- Web : www.genoscope.cns.fr)		
DEFINITION	Determination of this BAC-end sequence was carried out as part of a		
DEFINITION	collaboration with the Berkeley Drosophila Genome Project (BDGP).		
DEFINITION	The BDGP is constructing a physical map of the Drosophila		
DEFINITION	melanogaster genome using these BACs. For further information		
DEFINITION	please see http://www.fruitfly.org The BDGP Drosophila		
DEFINITION	melanogaster BAC library was prepared by Kazutoyo Osoegawa and		
DEFINITION	Aaron Mamoser in Pieter de Jong's laboratory in the Department of		
DEFINITION	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,		
DEFINITION	NY. The library is named RPCL-98 and was constructed by partial		
DEFINITION	EcoRI digestion of Drosophila DNA provided by the BDGP from the		
DEFINITION	isogenic strain Y2; cn bw sp, the same strain used for the BDGP's		
DEFINITION	P1 and EST libraries. A more detailed description of the library		
DEFINITION	and how to order individual BAC clones, the entire library, or		
DEFINITION	filters for hybridization from the BACPAC Resource Center can be		
DEFINITION	found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .		
DEFINITION	location/Qualifiers		
DEFINITION	1..1001		
DEFINITION	/organism="Drosophila melanogaster"		
DEFINITION	/mol_type="genomic DNA"		
DEFINITION	/db_xref="taxon:7227"		
DEFINITION	/clone="BACR13009"		
DEFINITION	/clone_lib="RPCL-98"		
DEFINITION	/note="end : T7"		
ORIGIN			
Query Match	11.5%;	Score 55.6;	DB 10; Length 1001;
Best Local Similarity	38.3%;	Pred. No. 0.027;	
Matches 157;	Conservative 42;	Mismatches 211;	Indels 0; Gaps 0
Qy	51	TATTTTACTGTGATACCTTACAAAGTGTATATAGTGTGTAACTTATTTATTTGAGA	110
Db	54	TATTTTCTACTATMYTTTTTWTWTTTTTATTTTTTWTWTTTTTATTTTTTWTWTTT	113
Qy	111	GGATTTTAAACACCTTGAAGCTAAACTTAAATAATAATTTCTATCTTTAAAG	170

QY	DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT	FEATURES	ORIGIN	QY
114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT	114	ATTWMTTTCATATTTT
171	GCACATATTTACGTCG	171	GCACATATTTACGTCG	171	GCACATATTTACGTCG	171	GCACATATTTACGTCG	171	GCACATATTTACGTCG	171	GCACATATTTACGTCG	171	GCACATATTTACGTCG	171	GCACATATTTACGTCG
174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC	174	ATTCACATTTTATTTAC
231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG	231	AAATCTCTTACACAGG
234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT	234	TTTTTCTCTTCACAT
291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA	291	CTGTGACATTTCCAA
294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT	294	TTTTTCTCTTCACAT
351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC	351	ACCGATGAACTTCTC
354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT	354	TTTTTCTCTTCACAT
411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG	411	ATCTAAGCAAGTTTG
414	TACTTWTTCATATTT	414	TACTTWTTCATATTT	414	TACTTWTTCATATTT	414	TACTTWTTCATATTT	414	TACTTWTTCATATTT	414	TACTTWTTCATATTT	414	TACTTWTTCATATTT	414	TACTTWTTCATATTT
CNS00396		CNS00396		CNS00396		CNS00396		CNS00396		CNS00396		CNS00396		CNS00396	
Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #		Drosophila melanogaster genome survey sequence TE13 end of BAC #	
BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		BACR08K10 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	
AL063921		AL063921		AL063921		AL063921		AL063921		AL063921		AL063921		AL063921	
GI:494178		GI:494178		GI:494178		GI:494178		GI:494178		GI:494178		GI:494178		GI:494178	
Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)		Drosophila melanogaster (fruit fly)	
Drosophila melanogaster		Drosophila melanogaster		Drosophila melanogaster		Drosophila melanogaster		Drosophila melanogaster		Drosophila melanogaster		Drosophila melanogaster		Drosophila melanogaster	
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
Genoscope.		Genoscope.		Genoscope.		Genoscope.		Genoscope.		Genoscope.		Genoscope.		Genoscope.	
Direct Submission		Direct Submission		Direct Submission		Direct Submission		Direct Submission		Direct Submission		Direct Submission		Direct Submission	
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 19													

```

Db      500  TWAADAWTTTTTTTTAAAAAABAAAAAABAAAAAABAAAAAATAAATATTTTTTTTTAAW 559
Qy      103  ATTGAGAGGATTTTAAACASACGCTAGAACSTAAACSTAAATAAATATTTCTCAT 162
Db      560  ATAAACMCAAAUUNTTTTUUNUUUUTUUTMTWNHNTUAMAHNTTUUNUUNUUAAMNH 619
Qy      163  CTTTAAAGSACATATTAACGTGGCTAAGSACATTAACAGCTATTAAGTAAACCTCATG 222
Db      620  MTWNHNTUAMHNTTTUATAAUUUUUTCMUUNUUNHNBABAAAMWTNTHTWTAUNHAT 679
Qy      223  TCGGCASTAAATTTCTTAAACAGCGTTCTGTCTTTTCAAAGGACTCCAAATATGSCA 282
Db      680  UHUUUUUCAMMCMCTHTCENHUUUNHUTAHNTHTNHUAAUUMUUYMUUUYMUUJSTAC 739
Qy      283  CTAATTATCTGTGGCACTTCCAAATTATATCTCCCTATUGGATTTATGTGGCCGTTT 342
Db      740  TUNHNNHNUAAUUTTUAYUAMAHAMMMHNNHNUAAAAAAMAAWTNTNHNTTNNHNTU 799
Qy      343  AAATAGTACCGAGTTGATCTTCACTGTGTGGAGTTGTCTTTGCTCTCTAAAGCTC 402
Db      800  UHUUUUTCTCMUCTNCHMUUNHATUCTMTNTNMHMTWTWNHNTHTNNHTTNNHMAWHT 859
Qy      403  TTTCAATTATCTAAAGCAAGTTTGCTATTAATCAADA 439
Db      860  WCMWMMHAAUUTWATHCHACMHTNNHNNHNNHNNHNNHNA 896

```

RESULT 4	CNS00EVL/c	LOCUS	DEFINITION
	CNS00EVL	1101 bp	DNA linear GSS 04-JUN-1999
			Drosophila melanogaster genome survey sequence T7 end of BAC:
			BACR29823 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION	AL069706
VERSION	AL069706.1
KEYWORDS	GI:4949849
SOURCE	GSS.
ORGANISM	Drosophila melanogaster (fruit fly)
REFERENCE	Drosophila melanogaster
TITLE	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
JOURNAL	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephyndroidea; Drosophilidae; Drosophila.
	1 (bases 1 to 1101)
AUTHORS	Genoscope.
	Direct Submission
	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage

COMMENT

BP 191 91006 EVRY ceex - FRANCE (E-mail : segreidgenoscope.cns.it)
- Web : www.genoscope.cns.it

Determination of this BAC-end and sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCT-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila.bac.htm>.

```
FEATURES
source
location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29523"
/clone_1fb="PCR-98"
/note="end : T7"
/note="end : T7"

```

[illegible]

RESULT 5	CL104745	LOCUS	DEFINITION
	CL104745	1191 bp	DNA
	ISB1-43112.T7.1	Xenopus tropicalis	genomic clone ISB1-43112,
			genomic survey sequence.

ACCESSION	CL104745
VERSION	CL104745.1
KEYWORDS	GI:40598380
SOURCE	GSS.
ORGANISM	Xenopus tropicalis (western clawed frog)
	Xenopus tropicalis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;

REFERENCE
1 (pages 1 to 1191)
Krementzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson

```

Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watsn.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 36
High quality sequence stop: 397.
Location/Qualifiers
1..1191
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-43112"
/clone_1b="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC

```

ORIGIN	Library Segment 1
Query Match	9.3%; Score 47.8; DB 10; Length 1191;
Best Local Similarity	46.2%; Prod. No. 1.6;

Matches 192; Conservative 0; Mismatches 223; Indels 1; Gaps 1;

43 TATGAAATATTTTACTGTGATCTTCAAGTGTATGATGCTGTGTAATCTATTT 102
 606 TATTTAAATTTTATTTATTTATTTTAAATTTATTTATTTATTTTATTTT 665
 103 ATTGAGAGGTATTTTAAACACCTTGAACCTTAAATTAATTAATAT-TTCTCTA 161
 666 TTAATATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 725
 162 TCTTTAAAGGCACATTTACGCTTAAAGGCATTTACGCTGATTTACGTTAAATCTAT 221
 726 TTTTAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 785
 222 GTGCGACATTAATTTCTTTACACGCGTCTCTCTTTCCAAAGGACCTCCAAATAGCC 281
 786 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 845
 282 ACTATTTATCTGTGCAATTTCCAAATTTATTTCCCTATTTGGGTAATTGATGTGCCGTT 341
 846 TTTATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 905
 342 TAAATAGTACGATTTGATCTTCACTGTGAGTTTGTCTTTGCTTCTCTAAAGGT 401
 906 TTTTATTTTATTTATTTTATTTTATTTATTTATTTATTTTATTTTATTTATTT 965
 402 CTTCATTTATCTTAAAGCAAGTTTGTATTAATTTCAAAATCTTGTCTTCTCCAT 457
 966 TTTTATTTTATTTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTT 1021

RESULT 6
 BH945927
 LOCUS BH945927 705 bp DNA linear GSS 01-OCT-2002
 DEFINITION chub8a02.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
 sequence.
 ACCESSION BH945927
 VERSION BH945927.1 GI:23425987
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 705)
 Delhaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.
 Whole genome shotgun reads from Brassica oleracea
 Unpublished (2002)
 CONTACT: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.edu
 Plate: chub8 row: a column: 02
 Seq primer: -210bpOT forward
 Class: shotgun
 High quality sequence start: 26
 High quality sequence stop: 551.
 Location/Qualifiers
 1..705
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea002"
 /note="Vector: pOTw13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea T01000H3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."

ORIGIN

Query Match 9.8%; Score 47.6; DB 9; Length 705;
 Best Local Similarity 45.5%; Pred. No. 1.8;
 Matches 170; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

53 TTTTACGATCTTCAAGTGTATGATGCTGTGTAATCTTATTTATTTGAGAG 112
 225 TTTTCACTTGATTTAGTCAAGAAAATTTAAAGTAAAGATTAACAAATTTTATTTA 284
 113 TATTTTAAACACCTTGAACCTTAAATTAATTAATTTCTTATCTTTAAAGGC 172
 285 TTTTAAAGTTATTTATTTAAATCTATTAATTAATTTATTTATTTATTTAATTA 344
 173 ACATTTACGCTTAAAGCAATTTACGCTGATTTATCTTTAACTCATGTGCCACTTA 232
 345 ATTCATCTCCAAAGTTAATCAATGAAAATTTAATTAATTAATTAATTAATTAAGAGT 404
 223 ATTCCTTACACGCGTCTGCTCTTTCCAAAGGACCTCCGAATATGCCACTATTTATCT 292
 405 TATTTTAAAGTATATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTAT 464
 293 GTGCAATTTCCAAATTTATTTCCCTATTTGGGTAATTGATGTGCCGTTTAAATAGTCAC 352
 465 AAGATTTTAAACAATTAATTTATTTAATATTTTATTTATTTATTTATTTATTTAATA 524
 353 CGATTGAATCTTCACTTGTGAGTTTGTCTTTGCTTCTCTTAAAGGTCTTCAATTTAT 412
 525 TGATTTTCTTTTAAATTTTAAAGTTTATCTTAATTAATTTGTAAATATTTTAAACATTT 584
 413 CTAAAGCAAGTTT 426
 585 GTAAACTTGTGTT 598

RESULT 7
 CNS0106X/c
 LOCUS CNS0106X/c 1101 bp DNA linear GSS 26-JUL-1999
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
 BACN03K20 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL098595
 VERSION AL098595.1 GI:5610206
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 Web : www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (BDGP) -
 http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billand at CBPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBeoBAC11.

FEATURES
 source
 1..1101
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone_lib="BACN03K20"
 /clone_lib="DrosBAC"
 /plasmid="pBeoBAC11"
 /note="end : T7"

ORIGIN

Query Match 9.7%; Score 47; DB 10; Length 1101;

	Query Match	9.5%	Score 46.2;	DB 1;	Length 1161;	
	Best Local Similarity	42.8%	Pred. No. 3	8;		
	Matches 171;	Conservative	0;	Mismatches	227;	Indels 2; Gaps 1;
Oy	51 TATTTTACTGTGATACCTTAACAAGTGATATATGGTGTGTGA	CTTATTATTTGAGA	110			
Db	756 TATTTTATGAGNGATAATATATATATTTTTATATTTTGA	AATATTTATTTANN	815			
Oy	111 GGATTTTAAACACCCTTAGA	CTTAAACCTTAATAATATTTCTCATCTTTAAAG	170			
Db	816 GATTAATTTTATATANTTTAAAANNTTAAATTTANN	TNNTTATATATTTTATATNNTTAA	875			
Oy	171 GCACATATTAACGTGGCTPAAGCAATPACAGCTGATATAC	TGTAACATGCATGTGCCACT	230			
Db	876 GTATATTTATATATTTATTTATTTATTTAGTTTNGAAT	TTCATTAATATTTATATATAT	935			
Oy	231 AAATCTTCTPAACACCGGTTCTGTCTTTCCAAAGGAC	TCCGAATATGCCACTATTTAT	290			
Db	936 TNATTTTATTAATATTAATTTACATTAATTTATTTT	AGNNTGTTTATATATATTTGA	995			
Oy	291 CTGCGCATTTCCATTTATATTCCTATTCGGATTTGA	TATGATGGCCGTTTAAATATGC	350			
Db	996 GTAAG-ATTTTATATTTTTTTTNNCNCNTGTTAT	TATATTTTTTTTTTTTAAATATTTT	1053			
Oy	351 ACCATTTGAATCTTCATCTGTGCGAGTTTGTCTTT	GCTTCTCTCTCTCTCTCTCTCAATTT	410			
Db	1054 CTTNTTTATTTNTTNNTATTTTTTTATTTGTTNA	TATATATTAANAATTTATATGATANTTA	1113			
Oy	411 ATCTAAGCAAGTTTGTATATATTCAAATATCTT	GCTTT 450				
Db	1114 ATTAATATATGATTAATGAATATGTTAANA	TTTTATTTT 1153				
RESULT 11						
LOCUS	AJ928310/c					
DEFINITION	AJ928310 Theileria annulata merozoite	Theileria annulata cDNA clone				
VERSION	AJ928310					
KEYWORDS	AJ928310.1 GI:67498696					
SOURCE	EST.					
ORGANISM	Theileria annulata					
REFERENCE	Theileria annulata					
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileridae;					
	Theileria.					
	1 (bases 1 to 1079)					
	Pain,A., Renaud,H., Berriman,M., Murphy,L., Yeats,C.A., Weir,W.,					
	Kethornou,A., Ajalett,M., Bishop,R., Buchler,C., Cochet,M.,					
	Coulson,R.M.R., Cronin,A., de Villiers,E., Fraser,A., Posker,N.,					
	Gardner,M., Gobie,A., Griffiths-Jones,S., Harris,D.E., Katzer,F.,					
	Lake,N., Lord,A., Maser,P., McKellar,S., Mooney,P., Morton,F.,					
	Nene,V., O'Neill,S., Price,C., Quail,M.A., Rabinowitsch,E.,					
	Raillings,N.D., Rutter,S., Saunders,J., Seeger,K., Shah,T.,					
	Squires,R., Squires,S., Tivey,A., Walker,A.R., Woodward,J.,					
	Doelaere,D.A.E., Langsley,G., Rajandream,M.-A., McKeever,D.,					
	Shiels,B., Tait,A., Barrell,B. and Hall,N.					
TITLE	The genome of the host-cell transforming parasite Theileria					
JOURNAL	annulata and a comparison with T. parva					
COMMENT	Unpublished (2005)					
	Contact: Pain A					
	The Pathogen Sequencing Unit					
	The Wellcome Trust Sanger Institute					
	Genome Campus, CB10 1SA, UNITED KINGDOM					
	Merozoite cDNA library; Frank Katzer and Brian Shiels, Division of					
	Veterinary Infection and Immunity, ICM, University of Glasgow, UK.					
FEATURES	Location/Qualifiers					
source	1..1079					
	/organism="Theileria annulata"					
	/mol_type="mRNA"					
	/isolate="Ankara (clone D7) "					
	/db_xref="taxon:5874"					

Query Match	9.5%	Score 46.2;	DB 1;	Length 1161;
Best Local Similarity	42.8%	Pred. No. 3	8;	
Matches 171;	Conservative	0;	Mismatches 227;	Indels 2; Gaps 1;
Dy	51	TATTTTACTGTGATACCTTAACAAGTGATATATGGTGTGTGA	CTTATTATTTGAGA	110
Db	756	TATTTTATGAGNGATAATATATATATTTTATATTTTGAATATATTTTANN		815
Oy	111	GGAATTTAAACACCCTTAGA	CTTAAACCTTAATAATATTTCTCATCTTTAAAG	170
Db	816	GATTAATTTTATATANTTTAAAANNTTAAATTTNNNTTATATTTTATNTNNTAA		875
Oy	171	GCAATATTTACGTGGCTPAAGCAATPACAGCTGATATCTGTAAACTCATGTGCACAT		230
Db	876	GTAATTTATTTATTTATTTATTTATTTAGTTTNGAATTTTGA	TGATATATTTATTTATTTAT	935
Oy	231	AAATCTTCTPAACACCGGTCGTCTCTTTCCAAAGCAGCTCGAATATGCCACTATTTAT		290
Db	936	TNATTTTATTTAATPAAGTTATTTACATTAATTTATTTTAA	NNNTTTTATATPATTTGA	995
Oy	291	CTGTGGCATTTCCATTTATATTCCTATTTGGGATTTATATGTGGCCGTTTAAATATGC		350
Db	996	GTAAG-ATTTTTATTTTTTTTTTNNTCNCTGTTATTTATTTTTTTTTTTTAAATTTTT		1053
Oy	351	ACCATTTGAATCTTCATCTGTGCGAGTTTGTCTTTGCTTCTCTCTCAAAGCTTTCAATTT		410
Db	1054	CTTTNTTTATTTNTTNTTATTTTTTTTATTTGTMTAATATATTTAANAATTTTATGATANTTTA		1113
Oy	411	ATCTAAGCAAGTTTGTATTAATTCAAATATCTTTGCTTT		450
Db	1114	ATTAATATATGATTAATGAATAATGTTAANAATTTTATTTT		1153
RESULT 11	AJ928310/c	1079 bp	mRNA	linear EST 10-JUN-2005
LOCUS	AJ928310			
DEFINITION	Theileria annulata mezoite Theileria annulata cDNA clone			
VERSION	AJ928310			
KEYWORDS	AJ928310.1 GI:67498696			
SOURCE	EST.			
ORGANISM	Theileria annulata			
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileridae;			
AUTHORS	Theileria.			
1 (bases 1 to 1079)				
Pain,A., Renaud,H., Berriman,M., Murphy,L., Yeats,C.A., Weir,W., Kethornou,A., Ajalett,M., Bishop,R., Buchler,C., Cochet,M., Coulson,R.M.R., Cronin,A., de Villiers,E., Fraser,A., Posker,N., Gardner,M., Gobie,A., Griffiths-Jones,S., Harris,D.E., Katzer,F., Lake,N., Lord,A., Maser,P., McKellar,S., Mooney,P., Morton,F., Nene,V., O'Neill,S., Price,C., Quail,M.A., Rabinowitsch,E., Rawlings,N.D., Rutter,S., Saunders,J., Seeger,K., Shah,T., Squares,R., Squares,S., Tivey,A., Walker,A.R., Woodward,J., Dohelaer,D.A.E., Langsley,G., Rajandream,M.-A., McKeever,D., Shiels,B., Tait,A., Barrell,B. and Hall,N.				
Db	111	The genome of the host-cell transforming parasite Theileria annulata and a comparison with T. parva unpublished (2005)		
JOURNAL COMMENT	Contact: Pain A			
FEATURES	The Pathogen Sequencing Unit The Wellcome Trust Sanger Institute Genome Campus, CB10 1SA, UNITED KINGDOM Mezozote cDNA library: Frank Katzer and Brian Shiels, Division of Veterinary Infection and Immunity, ICM, University of Glasgow, UK. Location/Qualifiers 1..1079 /organism="Theileria annulata" /mol_type="mRNA" /isolate="Ankara (clone D7)" /db_xref="taxon:5874"			

/organism="Populus deltoides"
/mol_type="mRNA"
/db_xref="taxon:3696"
/clone_lib="leaf inoculated with Marssonina pathogen of
Populus deltoides"
/note="Vector: PBK-CMV; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 9.3%; Score 45.2; DB 8; Length 698;

Best Local Similarity 46.6%; Pred. No. 6.4;
Matches 180; Conservative 0; Mismatches 203; Indels 3; Gaps 1;

QY 44 ATGAAATTTTCTGCTGATCTTACAGTTGATATAGTGTGTTAATTATTTA 103
DB 24 ATCAAAATTTGATTTTCTTCTGCTCAAAATTTTCTTCTTCTTCTTCTG 83
QY 104 TTGAGAGTATTTTACACACCTTACAACTTAATAATTAATTTCTCTATC 163
DB 84 TTTGATGCTGATATCAAAAATTTAAATTTAAATAAAACATTATTTAATG 143
QY 164 TTTAAAGCACATATTCGTGCTAGGCAATTAACAGTATATCTGTAACATC 223
DB 144 TTTTCAAGGAAAGAACTTTGAAAAGCAATTTTACACATTTCC--AAACAT 200
QY 224 GCCCAATAATTTCTTCTAAACGCGTCTCTCTTCCAAAGGACTCCGAAT 283
DB 201 CCTGATATATTTAACTGAGATTTATTTTCTTCTGAGTCTGATCATCAG 260
QY 284 TATTTATCTGTCGATTTCCATTTATTTATCCCTATTTGGTATTTGATG 343
DB 261 AAGATTAACATTTGTATTTGAAATTTAAAGTCTTTTGTGTTCTTCTGCA 320
QY 344 AATAGTACCGATTTGATCTTCTGCTGCTGCTTCTTCTTCTTCTTCTT 403
DB 321 ATTTGTTGACCGAGCAACTGTTTATTTAATTCATTTATTTCTGAAACGT 380
QY 404 TCAATTTATCTTAAAGCAAGTTTGTGA 429
DB 381 TTTGTTCTCTCTCAAAAGCTTCTTA 406

RESULT 14 718 bp DNA linear GSS 25-SEP-2003
CE216145 tigr-gss-dog-17000373053119 Dog Library Canis familiaris genomic,
LOCUS genomic survey sequence.
DEFINITION
ACCESSION CE216145
VERSION CE216145
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris

REFERENCE 1 (bases 1 to 718)
AUTHORS Kirschner,E.F., Balina,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
PUBMED 14512627
COMMENT Contact: Kirschner EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirchner@tigr.org
Class: shotgun.

FEATURES
source 1..718
location/Qualifiers
/organism="Canis familiaris"
/mol_type="genomic DNA"

/strain="Standard Poolle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BactXI; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 9.3%; Score 45.2; DB 9; Length 718;

Best Local Similarity 59.2%; Pred. No. 6.4;
Matches 77; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 51 TATTTTCTGCTGATCTTACAGTTGATATAGTGTGTTAATTATTTATTTGA 110
DB 418 TTTTATATTTGAGCTGATTTAGTTATTTATTAAGTAAATTTCTTTTCCAA 477
QY 111 GGTATTTTACACACCTTACAACTTAATAATTAATTTCTCTATTTAAG 170
DB 478 GCTATTTTGAACACAGATTTTCTTGAACATTAATCTATTTTATATACCA 537
QY 171 GCACATATTA 180
DB 538 TAACATATTA 547

RESULT 15 215 bp mRNA linear EST 29-JUN-1999
AV112242
LOCUS AV112242 Mus musculus C57BL/6J 10-day embryo Mus musculus cDNA
DEFINITION clone 2610010G18, mRNA sequence.
ACCESSION AV112242
VERSION AV112242.1 GI:5266322
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 215)
AUTHORS Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
Hara,A., Hayatsu,N., Horii,F., Ishikawa,T., Itoh,M., Izawa,M.,
Kawai,Y., Kikuchi,N., Kojima,Y., Matsumoto,T., Nishitani,H., Oda,H.,
Owa,C., Sato,K., Shibata,Y., Shigemoto,T., Shiraki,T., Sogabe,Y.,
Tomihara,Y., Suzuki,H., Tateo,M., Tomaru,Y.,
Tomihara,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs
JOURNAL Unpublished (1999)
COMMENT Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rcc.riken.go.jp
The mouse embryonic stem cell library and the activation of the embryonic
therapeutic and its application for the synthesis of full length cDNA
(Proc. Natl. Acad. Sci. U.S.A. 95(12):520-524 (1998))
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
Please visit our web site (http://genome.rcc.riken.go.jp) for
further details.

FEATURES
source 1..215
location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2610010G18"
/sex="mixed"
/dev_stage="10-day embryo"
/clone_lib="Mus musculus C57BL/6J 10-day embryo"

ORIGIN

Query Match 9.3%; Score 45; DB 1; Length 215;
 Best Local Similarity 53.8%; Pred. No. 6.8;
 Matches 93; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY	296	GCAATTC	CAATTA	TATATCC	CTATTG	GGATTT	GAATG	AGCCG	TTAA	TAGTCA	CGA	355
DB	41	GAATTTCC	TTATAT	CTCCG	AGATG	TTTGTG	CTGTC	CTTCA	CCCA	CACTCA		100
QY	356	TTGAATCT	CACTT	GTGAG	TTTGTG	CTTGTG	CTTCT	CTA	AAAGT	CTTCA	TTAT	415
DB	101	ATGACCTT	TTTACT	TGAA	TTTCTG	CTCA	ATGAC	ATTG	GTCC	ATTG	GAAT	160
QY	416	AAGCAAGT	TTTGTAT	TAATCA	AAATAC	TTGCTT	TC	CCATG	ACTTGA	ACT		468
DB	161	AGCCAAAT	GATGAT	TTTGA	TGTTG	ATGTTT	TA	CTTA	TAAAT	TTAG	ACT	213

Search completed: December 28, 2005, 11:15:57
 Job time : 3844 secs

Page Blank (uspto)


```

: TITLE OF INVENTION: INHIBIT A NEW CLASS OF ASPARTYL PROTEINASES
:
: FILE REFERENCE: DC-0134
:
: CURRENT APPLICATION NUMBER: US/10/030,808
:
: PRIOR FILING DATE: 2002-05-03
:
: PRIOR APPLICATION NUMBER: 60/143,355
:
: PRIOR FILING DATE: 1999-07-12
:
: NUMBER OF SEQ ID NOS: 4
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 1
:
: LENGTH: 705
:
: TYPE: DNA
:
: ORGANISM: Staphylococcus aureus
:
US-10-030-808-1

```

Query Match	8.5%	Score 41;	DB 3;	Length 705;
Best Local Similarity	44.9%;	Pred. No. 0.35;		
Matches 155;	Conservative	0;	Mismatches 190;	Indels 0;
			Gaps	0;

Oy	5	TATTTTACGTGATACCTTACAAAGTGATGATPANGTGGTGTAACCTAATTAATTTGAGA	110
	318	TATTTTACGTGATACCGATATCACTTCTTAATGTAGATGTGCGTTAATTAATTTTA	377
Oy	111	GGATTTTAAACACCTTAGAAGCTAAACCTAATAATAATATTTCTATCTTTAAG	170
Db	378	TTGATAGTTTCTCTCGTTAAGTATGATTAATCAGTAGCTTTTATCATTAATTAGAT	437
Oy	171	GCACATATATGATGGCTAAGGCAATATACAGTATATCTGTAACATCATGTGCACT	230
Db	438	GACCAACCATATATCTACTTTTATATTTGGGACATATATGGTTATAGTGACGTTTACT	497
Oy	231	AAATTTCTTAAACACGCGTTCTGTCTCTTTCCAGAGGACTCGAATATGACATTTAT	290
Db	498	AAATATGCACTTCTTTGTTTTTCCCTCTCAATTCACATATTTATGTCATTTATATAC	557
Oy	291	CTGTCGATTTCCAAATTAATATCCCTATTTGGGATTTGATGTGCGCGTTAAATATGC	350
Db	558	ATTTCATATGCTGGTTAGTTAGTCTTAATATACAGATATTTAAGCGGATTTAAACATTT	617
Oy	351	ACCGATGGAATCTTCACTTGTTCGAGATTTGTGCTTTTGCTTCTCT	395
Db	618	ACCCCTGTTCATTAATTAATTTCTCTTTTATCAATTTACT	662

```

US-09-949-016-15328
RESULT 9
US-09-949-016-15328
; Sequence 15328, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15328
; LENGTH: 111235
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME//KEY: misc feature
; LOCATION: (1) - (111235)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15328
```

Query Match 8.4%; Score 40.6; DB 3; Length 111235;

[illegible]

RESULT 10
US-09-949-016-13214/C

```

/ Patent No.6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949, 016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO. 13214
/ LENGTH: 113701
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-949-016-13214

```

Query Match	8.3%	Score 40.2;	DB 3;	Length 113701;
Best Local Similarity	53.5%	Pred. No. 1.7;		
Matches 84;	Conservative	0;	Mismatches 73;	Indels 0;
				Gaps 0;

QY	41	AGTATGAAAANTTTTACTGTGACTCAACAGTTGATATATNGTGTCGTA	CTAT	100
Db	20871	ATTATATAAGTTATTATATAGTATATATATATAAATTATATATATTAATATG		20812
QY	101	TTATTGGAGGTATTTAAACAACCTTAGACTAAACTTAAATAATATTCTCT		160
Db	20811	TATATATATACCTTAATTTGATGCATATATATATATATATATATATATATATAT		20752
QY	161	ATCCTTAAAGCACATATTTAGTGGCTPAAGCAATTA		197
Db	20751	ATATATATATATATATATATATAGATTAAGAAGGAATA		20715

```

RESULT 11
US-09-790-988-1
? Sequence 1, Application US/09790988
? Patent No. 6632935
? GENERAL INFORMATION:
? APPLICANT: SHIGEMOBU, SHUJI
? APPLICANT: WATANABE, HIDEMI
? APPLICANT: HATTORI, MASAHIRA
? APPLICANT: SAKAKI, YOSHITUKU
? TITLE OF INVENTION: GEMOME DNA OF BACTERIAL SYMBIONT OF APHIDS
? FILE REFERENCE: 081356/0159
? CURRENT APPLICATION NUMBER: US/09/790, 988
? CURRENT FILING DATE: 2001-02-23
? PRIOR APPLICATION NUMBER: JP2000-107160
? PRIOR FILING DATE: 2000-04-07
? NUMBER OF SEQ ID NOS: 7
? SOFTWARE: PatentIn Ver. 2.11
? SEQ ID NO: 1
? LENGTH: 640681

```



```

RESULT 15
US-09-949-016-14810/C
: Sequence 14810, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14810
: LENGTH: 174170
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(174170)
: OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14810

```


Db 241 AACAGCGCTTGTCTCTTCCAAAGGAGCTCCGAATATGCCACTATTTATCTGTGGCATT 300
Qy 301 TCCAAATTTATATCCCTTATTTGGGATTTTGTATGTGGCCGCTTTAAATAGTCACCGATTGAA 360
Db 301 TCCAAATTTATATCCCTTATTTGGGATTTTGTATGTGGCCGCTTTAAATAGTCACCGATTGAA 360
Qy 361 TCTTCACTTGTTCGAGTTTGTCTTTGCTTCTCTTAAAGCTCTCAATTTATCTTAAAGCA 420
Db 361 TCTTCACTTGTTCGAGTTTGTCTTTGCTTCTCTTAAAGCTCTCAATTTATCTTAAAGCA 420
Qy 421 AGTTTGTATTAATTCAAATTAATCTTGTCTTCTCCATGACTTGAACCTCCAAATGATGAG 480
Db 421 AGTTTGTATTAATTCAAATTAATCTTGTCTTCTCCATGACTTGAACCTCCAAATGATGAG 480
Qy 481 GTACC 485
Db 481 GTACC 485
RESULT 2
US-10-600-230-81
; Sequence 81, Application US/10600230
; Publication No. US20040092020A1
; GENERAL INFORMATION:
; APPLICANT: WALKINSON, Jack
; APPLICANT: MCBRIDE, Kevin
; APPLICANT: BERTALIN, Sean
; TITLE OF INVENTION: GENETIC CONSTRUCTS HAVING HETEROLOGOUS
; TITLE OF INVENTION: 3' POLYADENYLATION SIGNAL SEQUENCE MOTIFS THAT FUNCTION IN
; FILE REFERENCE: 0325.210
; CURRENT APPLICATION NUMBER: US/10/600,230
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: 60/390,529
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 4046
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-10-600-230-81
Query Match 97.9%; Score 474.8; DB 7; Length 4046;
Best Local Similarity 99.6%; Pred. No. 3.8e-101;
Matches 476; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 6 CGGAAGAGGAAAGTACCTCTCGTTGCTAGACATGATGAATATTTTACTGTGAT 65
Db 3546 CGGAAGAGGAAAGTACCTCTCGTTGCTAGACATGATGAATATTTTACTGTGAT 3605
Qy 66 ACTTACAAGTTGATATATAGTGTGTGTAACTTATTTATTTGAGAGTATTTTAAACAC 125
Db 3606 ACTTACAAGTTGATATATAGTGTGTGTAACTTATTTATTTGAGAGTATTTTAAACAC 3665
Qy 126 CTTAGAACTTAAACTTATATTAATTAATTTCTCTATCTTTAAAGGACATATTAGTGG 185
Db 3666 CTTAGAACTTAAACTTATATTAATTAATTTCTCTATCTTTAAAGGACATATTAGTGG 3725
Qy 186 CTAAAGCAATTAAGCTGATATATCTGTAAACTCANGTGGCACTAAATCTTCTAACAC 245
Db 3726 CTAAAGCAATTAAGCTGATATATCTGTAAACTCANGTGGCACTAAATCTTCTAACAC 3785
Qy 246 GCGTCTGTCTCTTTCCAAAGGACTCCGAATATGCCATTTATCTGTGGCATTTCCAA 305
Db 3786 GCGTCTGTCTCTTTCCAAAGGACTCCGAATATGCCATTTATCTGTGGCATTTCCAA 3845
Qy 306 TTTATATATCCCTTATTTGGGATTTTGTATGTGGCCGCTTTAAATAGTCACCGATTGATCTTC 365
Db 3846 TTTATATATCCCTTATTTGGGATTTTGTATGTGGCCGCTTTAAATAGTCACCGATTGATCTTC 3905
Qy 366 ACTTGTGAGATTTTGTCTTTGCTTCTTAAAGGCTTCAATTTATCTTAAAGCAAGTTT 425

Db 3906 ACTTGTGAGATTTTGTCTTTTGTCTCTTAAAGCTCTCAATTTATCTTAAAGCAAGTTT 3965
Qy 426 TGTATTAATTCAAATTAATCTTGTCTTTCTCCATGACTTGAACCTCCAAATGATGAGTGA 483
Db 3966 TGTATTAATTCAAATTAATCTTGTCTTTCTCCATGACTTGAACCTCCAAATGATGAGTGA 4023
RESULT 3
US-10-027-632-114994/c
; Sequence 114994, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 114994
; LENGTH: 3254
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-114994
Query Match 9.7%; Score 47; DB 5; Length 3254;
Best Local Similarity 55.1%; Pred. No. 1.1;
Matches 92; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
Qy 289 ATCTGTGCAATTCGAATTAATATTCCTTATTTGGGATTTGATGTGGCCGTTAAATAG 348
Db 1492 ATCTTAATTTAGTAATAGTATTTTCTTCTTGAATTTTAAATTTGGCTCTTTTAAAG 1433
Qy 349 TCACCGATTTGATCTTCACTTGTTCGAGTTTGTCTTGTCTCTTAAAGTCTTCAAT 408
Db 1432 TCTTATATTTCTAGTGAATTCCTCATCTTTTCTTAAATCTTAAATGTTATGAT 1373
Qy 409 TTTATCAAAGCAAGTTTGTATTAATTCAAATTAATCTTGTCTTCTCC 455
Db 1372 TTTATTAAGATCTTGTATGTTAATTTATTAATGATGTTTGTCTTC 1326
RESULT 4
US-10-027-632-114994/c
; Sequence 114994, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483

QY 78 ATATATGGTTGTGTA CTTATTTATTTGAGAGGTA TTTTAACACACCTTAGACTAA 137

[illegible]

```

RESULT 7
US-10-311-455-691
; Sequence 691, Application US/10311455
; Publication No. US20030143606a1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation Status of Cytosine Residues in DNA
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIORITY FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 691
; LENGTH: 5739
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-691

```

	Query Match	Similarity	Score	DB	Length
Beet	Local	45.9%	Pred. No. 2.1		
Matches	194	Conservative	0	Mismatches	228; Indels 1; Gaps 1
Qy	31 TTGCAGACAGATGAAAAATATTTTACGTCATATCTACAGTGAATATAGTGTG 90				
Db	729 TTGATTAATATTTTTTTATTTTATTTTCGTTTTATTTAAAGTTGTTAAATTTTGGT 788				
Qy	91 TGTACTTATTTATTTGAGAGGATTTTAAACACCTTAGACTAAACTTAATTAATAA 150				
Db	789 TTTTAGTTTGTAGTTTGTGAATGTTTAGAAAAAGTTAATTAAT-AGTTTTTGATGTT 847				
Qy	151 AATATTTCTCATCTTTAAAGGACATATTACGGGCTAAGCAATTACGCTATATCT 210				
Db	848 AATTTTAAATTAATTTTGTGGAATTTTTTTTATTAATTAATTAATTTGTATTTGTA 907				
Qy	211 GTAAACATAGTGGCACAATAATTTCTTCACAGCGCTGTCTCTTCCAGGCACT 270				

Db	908	TTATTTTTTTTTGAGATTTGTTTTATTAATAATTTTTTGTTTTTTTTTTTTTGTAGTTT	967
Qy	271	CCGAATATGCACATTTATCTGTGGGATTTCCAAATTATATCCCATTTGGGATTTGG	330
Db	968	TTTTATATAAATTTTTTTTGTTAGTATTTTTTTTGTCAGCTTTAAATAATGTTGTATG	1027
Qy	331	ATGTGGCGTTAAATAATGACACGATGAATCTTCACCTGTTCAGTTTGTCTTTTGGCT	390
Db	1028	TTAAGTTTTTGTATTAATTAATTTTTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTT	1087
Qy	391	TCTCTAAAGGCTTCATATTTATCTAAAGCAAGTTTGTATATCAAAATTACTTTGCTTT	450
Db	1088	TTTTTATTTGTTTTTTTTTAATTTTTTTTTTAATAATTTTATATAGTTTATGAGATTTTATTTT	1147
Qy	451	TCT	453
Db	1148	TAT	1150

```

RESULT 8
US-10-221-613-36
; Sequence 36, Application US/10221613
; Publication No. US20040029123A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/0221.613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 36
; LENGTH: 6092
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically derived genomic DNA (Homo sapiens)
US-10-221-613-36

```

Query Match	9.5%	Score 46.2	DB 7	Length 6092
Best Local Similarity	46.8%	Pred. No. 2.1		
Matches 181	Conservative	0	Mismatches 203	Indels 3
				Gaps 1
Qy	62	TGATACCTACAGATGATATATGTTGTGTGTAACTATATATTTGAGAGGATATTTTAAAC	121	
Db	4944	TGTTTTTTTAAAGTTGTTTTTTTTTTTTTTTATATTTTTTTTTTTTTTTTTTTTTTTTTTTT	5003	
Qy	122	ACACCTTGAACTAAACTTAAATTAATTAATTTCTTACTTTAAAGCACATATTAC	181	
Db	5004	ATTTTGTTTTTAATAAATTTGTGAATAATTTAGTTTTTATATATTAATTAAGGTTTGAAATAT	5063	
Qy	182	GTGGCTAAGCAATACAGCTGATATCTGTAAAACTCATGTCCGACATAATTTCTTA	241	
Db	5064	AAATTGATATATATTAATAATTTGTTATATTTTATTAATTTTGTTTTAAATAATTTTAA	5123	
Qy	242	ACAGCGCTTGTCTCTTTCCAGAGGACTCCGATATGCCATATTATCTGTGCATTT	301	
Db	5124	TGTTCTATTTTATATTAATTAATATGATATAATAATA---TATATTTTGTGTGAAT	5186	

Db 363 TATTATATATTTTGGTTTAAATTTTTCGTTTATTTTATATATATATATAT 422
QY 404 TCAATTTATCTAAGCAAGTTTGTATAT 433
Db 423 CGTTTATTTTACGTTATTTTCGTTTAT 452

RESULT 12

US-10-363-345A-1844/c
; Sequence 1844, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 1844
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 1844
US-10-363-345A-1844

Query Match 9.2%; Score 44.4; DB 8; Length 517;
Best Local Similarity 44.6%; Pred. No. 2.3; Matches 174; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 44 ATGAAATATTTTACTGTGATCTACAAAGTTGATATAGTGTGCGTAACCTATTTA 103
Db 455 ATTTATATTTTTCGTTAAATTTTATTTAGATATTTTATTTATTTTATTTTTCGTT 396
QY 104 TTGAGAGGATTTTAAACACACCTTAGAATAAATAATTAATTTCTATC 163
Db 395 TTTTCGTTATTTTATTTTATTTATATATATATATATATATATATATATATATAT 336
QY 164 TTTAAAGCAGATTTACGTGCTAAGCAATTAACGTGATATCTGTAAGATCATGT 223
Db 335 GTATATTTATTTATTTTCGTCGAGGATTTATGTTTATTTTATTTTTCGTT 276
QY 224 CGGCCTAAATCTTCTAAGAGCGTTCGTCCTTCCAGGAGCTCCGAATATGCCAC 283
Db 275 ATGTATTTATTTTATTTTTCGTTATTTTCGTTATTTTATTTTATTTTTCG 216
QY 284 TATTATCTGTGACATTTCAATTTATTTCCCTATTTGGATTTGATGTGCCGTTTA 343
Db 215 GATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAG 156
QY 344 AATAGCACCAGATTTGATCTTCTGTTGAGTTTGTCTTCTTCTTAAGGCTCT 403
Db 155 TATTATATTTTTCGTTTATTTTATTTTTCGTTTATTTTATTTTATTTTATTTAT 96
QY 404 TCAATTTATCTAAGCAAGTTTGTATAT 433
Db 95 CGTTTATTTTACGTTATTTTCGTTTAT 66

RESULT 13

US-10-363-483A-1843
; Sequence 1843, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain

; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 1843
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 1843
US-10-363-483A-1843

Query Match 9.2%; Score 44.4; DB 9; Length 517;
Best Local Similarity 44.6%; Pred. No. 2.3; Matches 174; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

QY 44 ATGAAATATTTTACTGTGATCTACAAAGTTGATATAGTGTGCGTAACCTATTTA 103
Db 63 ATTTATATTTTTCGTTAAATTTTATTTAGATATTTTATTTTATTTTATTTTTCGTT 122
QY 104 TTGAGAGGATTTTAAACACACCTTAGAATAAATAATTAATTTCTATC 163
Db 123 TTTTCGTTATTTTATTTTATTTTATTTATATATATATATATATATATATATATAT 182
QY 164 TTTAAAGCAGATTTACGTGCTAAGCAATTAACGTGATATCTGTAAGATCATGT 223
Db 183 GTATATTTATTTATTTTCGTCGAGGATTTATGTTTATTTTATTTTTCGTT 242
QY 224 CGGCCTAAATCTTCTAAGAGCGTTCGTCCTTCCAGGAGCTCCGAATATGCCAC 283
Db 243 ATGTATTTATTTTATTTTTCGTTATTTTCGTTATTTTATTTTATTTTTCG 302
QY 284 TATTATCTGTGACATTTCAATTTATTTCCCTATTTGGATTTGATGTGCCGTTTA 343
Db 303 GATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAG 362
QY 344 AATAGCACCAGATTTGATCTTCTGTTGAGTTTGTCTTCTTCTTAAGGCTCT 403
Db 363 TATTATATTTTTCGTTTAAATTTTTCGTTTATTTTATTTTATTTTATTTATTTAT 422
QY 404 TCAATTTATCTAAGCAAGTTTGTATAT 433
Db 423 CGTTTATTTTACGTTATTTTCGTTTAT 452

RESULT 14

US-10-363-483A-1844/c
; Sequence 1844, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 1844
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 1844
US-10-363-483A-1844

Query Match 9.2%; Score 44.4; DB 9; Length 517;
Best Local Similarity 44.6%; Pred. No. 2.3; Matches 174; Conservative 0; Mismatches 216; Indels 0; Gaps 0;

```
QY 44 ATGAAATATTTTACTGTGATCTTACAGTTGATATAGTGTGCTGTAACCTATTTA 103
DB 455 ATTTATATTTTTCGTAAATTTTATTTAGATATTTTATTTATTTTAAATTTTCGTT 396
QY 104 TTTGAGAGGATTTTAAACACCTTGAACCTAAACCTAAATTAATTAATTTCTATC 163
DB 395 TTTTCGTATTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTTTC 336
QY 164 TTTAAAGGACATTTATTCGTGCTAAGCAATTAAGCAATTAAGCAATTAAGCAAT 223
DB 335 GTATATATTTATTTATTTTCGCGTGAAGTATTTGTTTATATTCGTTTTTTTTCGT 276
QY 224 CCGCACTAAATCTTCTTAACAGCGTCTGCTCTTTCAAGGACTCCGAATATGCCAC 283
DB 275 ATGTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTTCG 216
QY 284 TATTTATCTGTGCAATTTCCATTTATTTCCCTATTTGGATTTGATGTGGCGTTA 343
DB 215 GATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTACG 156
QY 344 AATAGTACCGATTTGATCTTCACTTTGCGATTTTGTCTTTGCTCTTAAGGCT 403
DB 155 TATTTATATTTTTCGTTTATTTTATTTTTCGTTTTTATTTTATTTTATTTATTTAT 96
QY 404 TCAATTTATCTAAGCAAGTTTGTATTAAT 433
DB 95 CGTTTTTTTTTACGTTATTTTCGTTTAT 66
```

RESULT 15

```
US-09-925-065A-53227
; Sequence 53227, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53227
; LENGTH: 1803
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-53227
```

```
Query Match 9.0%; Score 43.8; DB 4; Length 1803;
Best Local Similarity 50.7%; Pred. No. 4.9; Mismatches 102; Indels 0; Gaps 0;
Matches 105; Conservative 0;
```

```
QY 48 AATAATTTTACTGTGATCTTACAGTTGATATAGTGTGCTGTAACCTATTATTG 107
DB 129 AGAATATGTATTTCTTACCAAGATGTTGGGTGATGTGTAATATCTTTAT 188
QY 108 AAGGATTTTAAACACCTTGAACCTAAACTTAATTAATTAATTTCTATCTTTA 167
DB 189 CTAGTCTTAAATATTTTATTTAAGAAAAATTTATGAAAGATTTTATGTTAACTCT 248
QY 168 AAGCAATATTTACGTGCTAAGGCAATTAAGCAATTAAGCAATTAAGCAATGTCGC 227
```

```
DB 249 GACTGACATTTTGTGATTTAAAGCAATTTGCTTTGTTGATAGCATATTATGGAAGAT 308
QY 228 ACTAAATTTCTTACACAGCGTCTGT 254
DB 309 ATTGAATAGCACATTTTGTGATCTTT 335
```

Search completed: December 28, 2005, 11:32:43
Job time: 825 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 09:24:36 ; Search time 293 Seconds

(without alignments)
867.750 Million cell updates/sec

Title: US-10-600-230-1

Perfect score: 485 1 gcgcgcgcgaaggaagaagt.....cctccaatgatgaggtacc 485

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4172979 seqs, 262114271 residues

Total number of hits satisfying chosen parameters: 8345958

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications NA New.*

1: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq.*

*, Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41.8	8.6	177175	7 US-11-121-086-79	Sequence 79, Appl
2	41.6	8.6	177175	7 US-11-121-086-79	Sequence 79, Appl
3	38.8	8.0	148220	7 US-11-121-086-62	Sequence 62, Appl
4	38.8	7.8	1303	7 US-10-750-185-50224	Sequence 50224, A
5	38	7.8	95223	7 US-11-117-187-188	Sequence 188, Appl
6	38	7.8	118996	7 US-11-121-086-84	Sequence 84, Appl
7	37.4	7.7	684	6 US-10-750-185-55780	Sequence 55780, A
8	37.4	7.7	171486	7 US-11-121-086-105	Sequence 105, Appl
9	37	7.6	1150	6 US-10-750-185-64497	Sequence 64497, A
10	37	7.6	207835	7 US-11-121-086-39	Sequence 39, Appl
11	37	7.6	207835	7 US-11-121-086-40	Sequence 40, Appl
12	36.6	7.5	201	6 US-10-995-561-73227	Sequence 73227, A
13	36.6	7.5	1048	6 US-10-750-185-30231	Sequence 30231, A
14	36.2	7.5	1588	6 US-10-750-185-39457	Sequence 39457, A
15	36.2	7.5	49979	6 US-10-995-561-13443	Sequence 13443, A
16	36	7.4	2248	6 US-10-750-185-52939	Sequence 52939, A
17	36	7.4	165156	6 US-10-995-561-13304	Sequence 13304, A
18	35.8	7.4	10467	6 US-10-240-708-2	Sequence 2, Appl1
19	35.8	7.4	645179	6 US-10-995-561-13293	Sequence 13293, A
20	35.6	7.3	3499	6 US-10-750-185-46381	Sequence 46381, A
21	35.6	7.3	156735	7 US-11-121-086-93	Sequence 93, Appl
22	35.4	7.3	1579	6 US-10-750-185-36438	Sequence 36438, A
23	35.4	7.3	19513	6 US-10-240-708-40	Sequence 40, Appl

C 24	35.2	7.3	3350	6 US-10-793-626-3453	Sequence 3453, Ap
C 25	35.2	7.3	151870	6 US-10-995-561-13199	Sequence 13199, A
C 26	35	7.2	887	6 US-10-750-185-52828	Sequence 52828, A
C 27	35	7.2	203467	7 US-11-121-086-50	Sequence 50, Appl
C 28	34.8	7.2	6866	6 US-10-240-708-20	Sequence 20, Appl
C 29	34.8	7.2	7673	6 US-10-516-768-15	Sequence 15, Appl
C 30	34.6	7.1	1654	6 US-10-750-185-50592	Sequence 50592, A
C 31	34.6	7.1	1825	6 US-10-750-185-58628	Sequence 58628, A
C 32	34.6	7.1	137935	6 US-10-995-561-13278	Sequence 13278, A
C 33	34.6	7.1	215308	7 US-11-121-086-77	Sequence 77, Appl
C 34	34.4	7.1	201	6 US-10-995-561-42981	Sequence 42981, A
C 35	34.4	7.1	2007	6 US-10-750-185-58341	Sequence 58341, A
C 36	34.4	7.1	1082144	7 US-11-117-187-211	Sequence 211, Appl
C 37	34.2	7.1	1755	6 US-10-750-185-24875	Sequence 24875, A
C 38	34.2	7.1	2629	6 US-10-276-231A-13	Sequence 13, Appl
C 39	34.2	7.1	403278	6 US-10-995-561-13421	Sequence 13421, A
C 40	34	7.0	1080	6 US-10-750-185-24697	Sequence 24697, A
C 41	34	7.0	2072	6 US-10-750-185-38146	Sequence 38146, A
C 42	34	7.0	2575	6 US-10-750-185-44880	Sequence 44880, A
C 43	34	7.0	1125000	6 US-10-995-561-13286	Sequence 13286, A
C 44	33.8	7.0	600	6 US-10-750-185-3073	Sequence 3073, Ap
C 45	33.8	7.0	600	6 US-10-750-185-3340	Sequence 3340, Ap

ALIGNMENTS

RESULT 1
US-11-121-086-79
; Sequence 79, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 79
; LENGTH: 177175
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-79

Query Match 8.6%; Score 41.8; DB 7; Length 177175;
Best Local Similarity 51.3%; Pred. No. 1.2;
Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 24 TCCTTCGTTGCGTACAGACAGTAAATTTTCTGATGATTAACAGTGAATAT 83
DB 76610 TCCTTCGATACATTAACATTTTCAAAACGCTGCTGTATGCTTAATATGATAT 76669
QY 84 GCTTCGTTGATCTATTTATTTGAGGATTTTAAACACCTTGAAGTAACTTAA 143
DB 76670 TATTTCACTGATTTATTTATTTCAATTAATTAATTAATTAATTAATTAAT 76729
QY 144 TAAATTAATTTCTTATCTTTTAAAGCATTATTCGCTTGAAGCAATTAACGCTG 203
DB 76730 ATTCTAGATCTTCATGATATATCTGTAACACTCTTTTAATCAATTAATTCCTCG 76789
QY 204 ATATACGT 212
DB 76790 TTTTATTT 76798

RESULT 2
US-10-240-708-62
; Sequence 62, Application US/10240708
; Publication No. US20050282157A1

```

; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/240,708
; PRIOR FILING DATE: 2002-10-03
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO: 62
; LENGTH: 6801
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; US-10-240-708-62

Query Match      8.6%; Score 41.6; DB 6; Length 6801;
Best Local Similarity 44.1%; Pred. No. 0.52;
Matches 173; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 60 TGTGATCTTACAGTGTGATATATGTTGTGTGTACTTATTTTGAAGAGTATTTA 119
    |||||
DB 5358 TTTGTAAATTAAGATTTTATTAAGGTTTATTAAGATTTATTAATTAATTTTNG 5417

QY 120 ACAACCTTAAGACCTTAATTAATTAATTTCTTAACCTTAAGACATAT 179
    |||||
DB 5418 TTTAATTAATGAAGATTAATTTTATTAATTAATTAAGATTAAGATTTAATTTT 5477

QY 180 ACGTGCTTAAGGCAATTAAGCTGATATACGTAAACTGATGTCCTCAATTTCTTC 239
    |||||
DB 5478 TCGAAGAGATGATGATGATGATGATTAATTAATTAATTAATTAATTTATTTA 5537

QY 240 TAAACAGCGTCTGTCTCTTCCAAAGGACTCCGAATATGCCATTTATCTGCGCAT 299
    |||||
DB 5538 GGGAAATGTAATTAAGATTAAGTATTAATTTATTTTATTTTATTTTATTTT 5597

QY 300 TTCCATTTATATTCCTTATGCGTATTTGATGCGCGTTAAATAGTACCGATTTGA 359
    |||||
DB 5598 TTGTAATTTAATTTTGTATTAATTTTATTTTATTTTATTTTATTTTATTTT 5657

QY 360 ATCTCACTTGTTCGATTTTGTCTTTGCTTCTTAAGGCTTCAATTTATCTAAGC 419
    |||||
DB 5658 TTTTATTTTATTTTGTGTTTGTGTTTATTTTATTTTATTTTATTTTATTTT 5717

QY 420 AAGTTTGTATTAATTCAAATACTTGTCTTT 451
    |||||
DB 5718 TTTTATTTTATTTATTTATTTGATTTGATTT 5749

RESULT 3
US-11-121-086-90/c
; Sequence 90, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
```

```

; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 90
; LENGTH: 148220
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-121-086-90

Query Match      8.0%; Score 38.8; DB 7; Length 148220;
Best Local Similarity 52.5%; Pred. No. 6.1;
Matches 85; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 17 AAGTACTCTTCGTTGCGTGAACAGTATGAATAATTTTACTGATATCTTACAGTT 76
    |||||
DB 77435 AATCATTTCTTTATGCGTGAATAGATTTATATATATATATATATATATATATA 77376

QY 77 GATTAATGTTGTGTGTACTTATTTTATTTGAGGATTTTAAACACCTTGAACCTAA 136
    |||||
DB 77375 TATATATATATATATACCAATATATATATATATATATATATATATATATATAC 77316

QY 137 AACTTAATAATAATATTTCTTATCTTTAAAGGACATAT 178
    |||||
DB 77315 ACACACACACACATATATATATATATATATATATAT 77274

RESULT 4
US-10-750-185-50224
; Sequence 50224, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50224
; LENGTH: 1303
; TYPE: DNA
; ORGANISM: Bovine
; US-10-750-185-50224

Query Match      7.8%; Score 38; DB 6; Length 1303;
Best Local Similarity 57.6%; Pred. No. 2.3;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 351 ACCGATTAATCTTCACTTGTTCGATTTTGTCTTTGCTTCTTAAGGCTTCAATTT 410
    |||||
DB 741 ACTTATTTCTGTTCATTTAGCTTTGTCTTCATTTCTTTTGTCTTCTTTTTC 800

QY 411 ATCTAAGCAAGTTTGTATTAATTCAAATFACCTTTGCTTTCTCATGACTGAACCT 468
    |||||
DB 801 AATTAATTAATTTTATTAATTCATCTTATCTCTTTTATTAATTCATTCACCAT 858

RESULT 5
US-11-117-187-188/c
; Sequence 188, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPEHAEVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
```

```

; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 188
; LENGTH: 95223
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (2720)..(72695)
; OTHER INFORMATION: N = A, C G, or T/U
US-11-117-187-188

```

```

Query Match
Best Local Similarity 7.8%; Score 38; DB 7; Length 95223;
Matches 92; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

```

```

QY 36 TAGACAGTATGAAATATTTTACTGTGATCTTGAAGTTGATTAATGTTGTGTGA 95
DB 24588 TATACATATTCATAAATAGTGTATCTTGGCTGACACATCTAAATGTGAATCTGAATA 24529
QY 96 CTTATTTATTTAGAGTATTTTAAACACACCTTAGAATAAATTAATAATTAAT 155
DB 24528 CACAGTACTATATTCATATGTTTACAGATTAATCTTAAACCAATTAAGATTTAT 24469
QY 156 TCTATCTTTTAAAGGCACATATTCGTGCTAAGGCAATTAACAGTGTATATCTGTA 215
DB 24468 TTTGTATTTTATATTAAGTTTCTGACTAAATAAATAATATCATGTAGTGTACC 24409
QY 216 AC 217
DB 24408 AC 24407

```

```

RESULT 6
US-11-121-086-84
; Sequence 84, Application US/11/121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 84
; LENGTH: 118996
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-84

```

```

Query Match
Best Local Similarity 7.8%; Score 38; DB 7; Length 118996;
Matches 131; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 51 TATTTTACTGTGATCTTACAGTTGATATATGTTGTGTAACTTATTTATTTGAGA 110
DB 9710 TCTTTCTATATTCATCTCATTTCTTAAAGTATGCTGTGGGTTATTTTACCCTTGTA 9769
QY 111 GGTATTTTAAACACCTTGAATAAATCTTAAATAAATAATTTCTTATCTTTAAAG 170
DB 9770 GCCGTTTCAATCTTTTCTGAAATATAACAAGATACATTAAGGTGTAATAATGATCTCT 9829

```

```

QY 171 GCACATATTTAGTGCTTAAAGCAATTAACGTGATATATCTGTAAACTCATGTCCACT 230
DB 9830 GTTTATATTTCCATCTACTCATCATGAACCTGAAGCTAGCGCTAATTAATCAATG 9889
QY 231 AAATCTCTTAAACAGCGTGTCTCTTTTCCAGGAGCTCCGATATGCCACTTTAT 290
DB 9890 ATGTCTTCATTAAGACCTTTCAGTCAATGAAGTGTGTCCCATGATTTCAATTAAGG 9949
QY 291 CTGTGCAATTTCCATTTATATTTCCCTATTTGGGATTTGATGTGG 336
DB 9950 CTGAATAAATCTTCATTTGCTGTAGTACCTCATAGCATGATGTGTGG 9995

```

```

RESULT 7
US-10-750-185-55780
; Sequence 55780, Application US/10/750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55780
; LENGTH: 684
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-55780

```

```

Query Match
Best Local Similarity 7.7%; Score 37.4; DB 6; Length 684;
Matches 77; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

```

```

QY 54 TTTTACTGTGATCTTACAGTTGATATATGTTGTGTAACTTATTTATTTGAGAGT 113
DB 504 TTTTCAATATATGATATTAAGGAAAGACATGTAATTTAAATTAATTTTGAAGACT 563
QY 114 ATTTTAAACACCTTGAATAAATCTTAAATAAATAATTTCTTATCTTTAAAGCA 173
DB 564 TTGCTATCAGCCTTTAAATGGAATGCATATATATATATATATAATTAATTAATAA 623
QY 174 CATATTAGTGGCTAAGCAATT 196
DB 624 TTTATATAATTTAAATATAATT 646

```

```

RESULT 8
US-11-121-086-105
; Sequence 105, Application US/11/121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 105
; LENGTH: 171486

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-105

```

Query Match	7.7%	Score 37.4	DB 7	Length 171486
Best Local Similarity	59.0%	Pred. No. 14		
Matches 82	Conservative 0	Mismatches 56	Indels 1	Gaps 1

Qy	Db	Qy	Db
43	17286	102	17346
TATGAAATATTTTCTGATGACCTTACAGTGTATATGGTGTGCTGA-ACCTATT	TATGAAATATTTTCTGATGACCTTACAGTGTATATGGTGTGCTGA-ACCTATT	TATTTGGAGGTATTTTACACACTAGAACTTAAATTAATTAATTTCTCTA	TATTTAGATATATATTAATTAATTTATATTTATATATATATTAATTAATTAATA
17345	17405	161	17405
TATGAAATATTTTCTGATGACCTTACAGTGTATATGGTGTGCTGA-ACCTATT	TATGAAATATTTTCTGATGACCTTACAGTGTATATGGTGTGCTGA-ACCTATT	TATTTGGAGGTATTTTACACACTAGAACTTAAATTAATTAATTTCTCTA	TATTTAGATATATATTAATTAATTTATATTTATATATATATTAATTAATTAATA

QY	162	TCTTAAAGGCACATATTA	180
Db	17406	TTTATAGATTATATATTA	17424

```

US-10-750-185-64497
US-10-750-185-64497
Sequence 64497, Application US/10750185
Publication NO. US20050260603A1
GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING
FILE REFERENCE: MM1100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIN version 3.1
SEQ ID NO 64497
LENGTH: 1150
TYPE: DNA
ORGANISM: Bovine 19866881242828
US-10-750-185-64497

```

Query Match	7.6%	Score 37	DB 6	Length 1150
Best Local Similarity	58.7%	Pred. No. 3.9	45	Indels 0
Matches 64	Conservative 0	Mismatches 45	Gaps 0	
Qy	32	TCGGTAGACAGTANGAAATATTTTACTGTACTTCAAGTTGATATATGTTGTGT	91	
Db	266	TACATTTAAATGACATATTTAATTAATAGATATCAAAATTCATTATCTTTGTTT	325	
Qy	92	GTAACATATTTATTTGAGAGTATTTTAAACACCCCTTAGACTTAAACT	140	
Db	326	ATAATTTGTTTTCTGAAGATATTTTGAAGGAGTATCAGAAAAAATAT	374	

RESULT 10
US-11-121-086-39
; Sequence 39, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04

```

; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 39
; LENGTH: 207835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-39

```

Query Match	7.6%	Score 37	DB 7	Length 207835
Best Local Similarity	48.4%	Pred. No. 18		
Matches 103	Conservative 0	Mismatches 110	Indels 0	Gaps 0

Accession	Gene	Length (bp)
QY 78	ATATATGCTTGCTGAACCTTATTTATGAGAGTATTTAACACACCTTACAAATACAA	159
Db 199888	ATGTTAAATTTAAATTTAAATTTAAATTTAGATACATTTTAAATATGTCAAGTCTCTCTTAA	199947
QY 138	ACTTAAATAAATTAATTTCTCTATCTTTAAAGCACAATTTACGTGCTAAAGCAATTA	197
Db 199948	GCTTCACAGAAATTAATATCAATCTCTTTAAATGAACAATTTATGATGAATTCGTCAAAA	200007

258 TTTCAGGACTCGAATATGCCATATTAT 290
200068 ATATTATATATGTGTAACTATTATCCAGACATTTCAATTTACATCTTGTGTCTG 200066

```

RESULT 11
US-11-121-086-40
; Sequence 40, Application US/1121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIORITY FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ. ID NO. 40
; LENGTH: 207835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-40

```

Query Match	7.6%;	Score 37;	DB 7;	Length 20/835;
Best Local Similarity	48.4%;	Pred. No. 18;		
Matches 103; Conservative	0;	Mismatches 110;	Indels 0;	Gaps 0;

QY 76 AATAATGGTGGGAGCAATATTAATGAGAGGATTTGATGCTGCTCTTAA 199947

Db 199888 ATGTAAAAATTTATAAAAATTTAAATTTAGATACATTAAAAATGTCAAGTCTCTTTA 199947

QY 138 ACTAATTAATATAATATTTCTCATCTTTAAAAAGCAGATATTAAGTGCGCTAAGGCATTA 197

Db 199948 GCTTCAGAAATATTAATCAATCTCTTTAAAAATGAACAAATTTATGATGAATTCGTCAAA 200067

QY 198 CACCTGATATACTGTAAAACTCATGTGCGCACTAAATCTTCTAACAGCGTTCTGTCTC 257

Db 200068 AATATTATATGGTTAACTCATTTACCAAGACATTTCAAAATTTACATCTTGTGTCTG 200067

QY	258	TTTCCAGGCACTCCGATATGCCACTATTAT	290
Db	200068	TTAAGACAGATTAGAAATACACTAATTTTAT	2001000

RESULT 12
US-10-995-561-73227/c

```

; NUMBER OF SEQ ID NOS: 107
;
; SOFTWARE: PatentIn version 3.3
;
; SEQ ID NO 39
;
; LENGTH: 207835
;

```

TYPE: DNA
ORGANISM: Homo sapiens
S-11-121-086-39

[illegible]

DB 199948 GCTTCAGCAATATTAATCATCTCTTAATAATGACCAATTTATGATGATTCGTCAAA 200007

258 TTTCAGGACTCGAATATGCCATATTAT 290
200068 ATATTATATATGTGTTAACTATTTACCAAGACATTTCAATTTACATCTTGTGCTG 200066

```

RESULT 11
US-11-121-086-40
; Sequence 40, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121, 086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567, 570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO. 40
; LENGTH: 207635
; TYPE: DNA

```

!	ORGANISM: Homo sapiens
US-11-121-086-40	
Query Match	7.6%; Score 37; DB 7; length 207835;
Best Local Similarity	48.4%; Pred. No. 18; Mismatches 0; Gaps 0;
Matches 103; Conservative	0; Indels 110; Gaps 0;
Qy	78 ATATANGTGTGGTGAACCTATTATTATTTGAGAGGATTTTAACAACCTTAGAACTAAA 137
Db	199888 ATGTAAATAATTTTAAATAATTTTAAATTTTAAATTTTGAATACATTTTAAATGTCAAAGTCTCCTTTA 199947
Qy	138 ACTAATTAATAATAAATATTTCTCATCTTTAAAGGCACATATTTACGTGGCTAAAGGCATTTA 197

[illegible]

Query Match	Best Local Similarity	Score	No. of Matches	Conservative	Mismatches	Indels	Gaps
Qy	7.5%	36.6	DB 6	Length 1048			
Db	49.2%	Pred. No. 4.8					
Qy	96	Conservative	0	Mismatches	99	Indels	0
Db	96	Conservative	0	Mismatches	99	Indels	0
Qy	283	CGATTATCTGCGGCACTTCCAACTTATATCCCGATGGGAGATTGANGGCCGCTT	342				
Db	60	CTCGCTCTTAGGCGCACTCCCTAGTATCTTAACTTATCATCTTAAATTAATGAA	119				
Qy	343	AAATAGTCAACGATTAATCTTCACTTGTGAGTTTTGTCTTTGCTCTTAAAGTTC	402				
Db	120	ACTCATTTGCTCAATTTAAATTAATCACTTGGCTTTTGAACAATTTTTCATTAACATTA	179				

```

: Sequence 13443, Application US/10995561
: Publication No. US20050272054A1
:
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
:
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 13443
:   LENGTH: 49979
:   TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(49979)
:

```

OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-10-995-561-13443)

Query Match 7.5%; Score 36.2; DB 6; Length 49979;

Best Local Similarity 50.3%; Pred. No. 19; Matches 86; Conservative 1; Mismatches 84; Indels 0; Gaps 0;

Qy 9 AAGAGGAAAGTGAAGTCTTGGGTAAGACATATGAAATATTTTACTGTGATCT 68
Db 43936 AACCAAGAGAAATACATTTATATCTAAACAATAGTCTTATGATAAAGGTTAGT 43937
Qy 69 TACAAGTATATATGTTGTGTGTAAGTATTTATTTAGAGGTATTTTAACACCTT 128
Db 43936 TTAAATGGATACAAAATTTGCTGTGTAATAAATGTTTCAAAATACATTCTATAGGTA 43877
Qy 129 AGAACTAAAACTTAATAATAATATTTCTATCTTTAAAGGACATATT 179
Db 43876 GAGACTATGTCTTAGTAAAGACAGTTATCTATTATCAAAATATCTATT 43826

Search completed: December 28, 2005, 11:37:47
Job time : 298 secs

November 2005

Published_Applications Nucleic Acid and Published_Applications Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

- Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).
- Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

This Page Blank (uspto)